



JOURNAL Patent: WO 0229038-A 1 11-APR-2002;

## FEATURES

FEATURES	Location/Qualifiers
source	1. .2881

source	1. .2881	/organism
--------	----------	-----------

```

/mol_type="unassigned DNA"

```

```

ORIGIN
/db_xref="taxon:9606"

```

STAT

```

Query Match      100.0%; Score 2881; DB 6; Length 2881;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

[illegible]



Db 781 TCGACGTGGACCGAGTGGTCCGCTCTGCAGCGCCAGCTGTGGGCGCGCTGGCGAARACGG 840  
QY 882 AGCCGAGCTGACCAACCCGCGCGCTCTCAACGCGGGCGCCTTTCTGTGAGGGCGAGAAT 941  
Db 841 AGCCGAGCTGACCAACCCGCGCGCTCTCAACGCGGGCGCCTTTCTGTGAGGGCGAGAAT 900  
QY 942 GTCCATGACCGACCGCTCTCTCTGCTGTGTCTCTGTGTGACCGGACGCTGGACCCGCTGG 1001  
Db 901 GTCCAGAA---AAGCCTGGCGCCACCCCTGTGCCCCAGTAGACGGCAGCTGGACCCGCTGG 957  
QY 1002 AGCAAGTGGTTCGGCCCTGTGGGCTGGAGCTGCACCCACTGGCGGAGCCGCTGAGTCTCTGAC 1061  
Db 958 AGCAAGTGGTTCGGCCCTGTGGGCTGGAGCTGCACCCACTGGCGGAGCCGCTGAGTCTCTGAC 1017  
QY 1062 CAGACACCCCGCAACCGAGGGGAGGAGTGCCAGGGGCACTGACCTGGACACCCGCAACTGT 1121  
Db 1018 CAGACACCCCGCAACCGAGGGGAGGAGTGCCAGGGCACTGACCTGGACACCCGCAACTGT 1077  
QY 1122 ACCAGTGACTCTGTGTACACAGTGTCTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1181  
Db 1078 ACCAGTGACTCTGTGTACACAGTGTCTCTGGCCCTGAGGACGTGGCCCTCTATGTGGGC 1137  
QY 1182 CTCAATCCCGTGGCCCTGTCCCTGGTCTCTGCTGCTCTGCTCTCATCTCTCGTTTATTGC 1241  
Db 1138 CTCAATCCCGTGGCCCTGTCCCTGGTCTCTGCTGCTCTGCTCTCATCTCTCGTTTATTGC 1197  
QY 1242 CGGAGAGAGGGGCTGACTCAGATGTGGCTGACTGTGTCMACTCTCACTCAGGCTTC 1301  
Db 1198 CGGAGAGAGGGGCTGACTCAGATGTGGCTGACTGTGTCMACTCTCACTCAGGCTTC 1257  
QY 1302 CAGCCGCTCAGCATCAAGCCCAAGCAAGCAGACAACCCCATCTGCTCAACCTCCAGCCG 1361  
Db 1258 CAGCCGCTCAGCATCAAGCCCAAGCAAGCAGACAACCCCATCTGCTCAACCTCCAGCCG 1317  
QY 1362 GACCTCAG---CACCAACCACTACAGGGCAGTCTCTGTCCCGCGGAGGATGGCCCC 1418  
Db 1318 GACCTCAGCAACCACTACAGGGCAGTCTCTGTCCCGCGGAGGATGGCCCC 1377  
QY 1419 AGCCCAAGTTCCAGTCAACCATGGGACCTGCTCAGCCCTTGGGTGGCGCCGCCAC 1478  
Db 1378 AGCCCAAGTTCCAGTCAACCATGGGACCTGCTCAGCCCTTGGGTGGCGCCGCCAC 1437  
QY 1479 ACACTGACACACAGCTCTCCACCTCTGAGGCGGAGGAGTCTGCTCCCGCCCTCTCCACC 1538  
Db 1438 ACACTGACACACAGCTCTCCACCTCTGAGGCGGAGGAGTCTGCTCCCGCCCTCTCCACC 1497  
QY 1539 CAGAACTACTTCGGTCTCCCTGCCGAGGCACCAACATGACCTATGGGACCTTCAAC 1598  
Db 1498 CAGAACTACTTCGGTCTCCCTGCCGAGGCACCAACATGACCTATGGGACCTTCAAC 1557  
QY 1599 TTCTTGGGGCGGGCTGATGATCCCTAATACAGGTATCAGCTCTCTATCCGCCAGAT 1658  
Db 1558 TTCTTGGGGCGGGCTGATGATCCCTAATACAGGTATCAGCTCTCTATCCGCCAGAT 1617  
QY 1659 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGAAGACGTG 1718  
Db 1618 GCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCCGAAGACGTG 1677  
QY 1719 AGTTGGCCCTAGTGGTGTGAGACCCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT 1778  
Db 1678 AGTTGGCCCTAGTGGTGTGAGACCCCTGCTGAGTCCCATCGTTAGCTGTGGACCCCT 1737  
QY 1779 GCGTCTCTCTCACCCGCGAGTCTCTGGCTATGACCACTGTTGGGAGCCGACCT 1838  
Db 1738 GCGTCTCTCTCACCCGCGAGTCTCTGGCTATGACCACTGTTGGGAGCCGACCT 1797  
QY 1839 GACAGTGGAGCTCGGCTCAAAAGCAGTCTGTCGAGGGCAGCTGGGAGCAGGATGTG 1898  
Db 1798 GACAGTGGAGCTCGGCTCAAAAGCAGTCTGTCGAGGGCAGCTGGG---AGATGTG 1854  
QY 1899 CTGCACTGGGGAGGAGCGCCCTCCACCTCTACTGTCAGCTGGAGGCCAGTGGC 1958

Db 1855 CTGCACCTGGCGAGGAGGGCCCTCCACCTCTACTACTGCCAGCTGGAGGCCAGTGCC 1914  
QY 1959 TGTACGTCTTTCACGAGAGCTGGCCCGCTTTTGCCCTGTGGAGAGGCCCTCAGCGTG 2018  
Db 1915 TGTACGTCTTTCACGAGAGCTGGCCCGCTTTTGCCCTGTGGAGAGGCCCTCAGCGTG 1974  
QY 2019 GCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTTGGCGCGGTGGCCTGCACCTCCCTCGAG 2078  
Db 1975 GCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTTGGCGCGGTGGCCTGCACCTCCCTCGAG 2034  
QY 2079 TACAAATCCCGGTCTACTGCTCATGACACCCAGATGCACTCAAGAGGTGGTGCAG 2138  
Db 2035 TACAAATCCCGGTCTACTGCTCATGACACCCAGATGCACTCAAGAGGTGGTGCAG 2094  
QY 2139 CTGGAGAGCAGCTGGGGGACAGCTGATCCAGGACCCAGCGGTCTCTGCACTTCAAAGAC 2198  
Db 2095 CTGGAGAGCAGCTGGGGGACAGCTGATCCAGGACCCAGCGGTCTCTGCACTTCAAAGAC 2154  
QY 2199 AGTTACCAAACTGCGCCTATCCATCCAGATGTGCCAGCTTCCCTGTGGAGAGTAAAG 2258  
Db 2155 AGTTACCAAACTGCGCCTATCCATCCAGATGTGCCAGCTTCCCTGTGGAGAGTAAAG 2214  
QY 2259 CTCTTGTCTCAGCTTACCAGAGATCCCTTTTATCACTCTGGAATGGCACGCGAGCGGTAC 2318  
Db 2215 CTCTTGTCTCAGCTTACCAGAGATCCCTTTTATCACTCTGGAATGGCACGCGAGCGGTAC 2274  
QY 2319 TTGCACTGCACTTCACTTCCAGCGGTGTACGCCAGCACTAGTACCTTGGCCTCAAG 2378  
Db 2275 TTGCACTGCACTTCACTTCCAGCGGTGTACGCCAGCACTAGTACCTTGGCCTCAAG 2334  
QY 2379 CTGTGGGTGTGAGGTGGAGGGCGACGGCAGAGCTTCAAGATCAACTCAATCACTCACC 2438  
Db 2335 CTGTGGGTGTGAGGTGGAGGGCGACGGGCGAGCTTCAAGATCAACTCAATCACTCACC 2394  
QY 2439 AAGGACACAAGCTTTGCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTCCCGAGCCCTG 2498  
Db 2395 AAGGACACAAGCTTTGCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGGTCCCGAGCCCTG 2454  
QY 2499 GTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAAAGATAATTTCCAGCCTG 2558  
Db 2455 GTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTGGCAGAAAGATAATTTCCAGCCTG 2514  
QY 2559 GACCCACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCGAGAACTCCACCTTGGAC 2618  
Db 2515 GACCCACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCGAGAACTCCACCTTGGAC 2574  
QY 2619 AGCCATCTCAGCTTCTTTGCTTCCCAAGCCCGACAGCCATGATCTCAACCTGTGG 2678  
Db 2575 AGCCATCTCAGCTTCTTTGCTTCCCAAGCCCGACAGCCATGATCTCAACCTGTGG 2634  
QY 2679 GAGGCGGGCAGCTTCCCNAAACGGGAACTCAGCCAGCTGGCTGCGAGCAGTGGCTGAGCTG 2738  
Db 2635 GAGGCGGGCAGCTTCCCNAAACGGGAACTCAGCCAGCTGGCTGCGAGCAGTGGCTGAGCTG 2694  
QY 2739 GCGCAGCCAGACGCTGGCTTTCAGAGTGTGGAGGCTGAGTGTGAGGCGCGCCAG 2796  
Db 2695 GCGCAGCCAGACGCTGGCTTTCAGAGTGTGGAGGCTGAGTGTGAGGCGCGCCAG 2752

RESULT 3  
AX451652  
LOCUS  
DEFINITION  
Sequence 1 from Patent WO0233080.  
ACCESSION  
AX451652  
VERSION  
AX451652.1  
GI:21698587  
KEYWORDS  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1  
Koehler, R.H.  
AUTHORS  
TITLE  
Regulation of human netrin binding membrane receptor unc5h-1



[illegible]



Db 730 GCGAGCTGAGCCCGTGGAGCAAGTGGTGGGCTGGAGTGCACCCACTGGCG 789

Qy 1044 AGCCGTGAGTGTCTGACACACACCCCGAACGAGGGGAGAGTGCAGGGCACTGAC 1103

Db 790 AGCCGTGAGTGTCTGACACACACCCCGAACGAGGGGAGAGTGCAGGGCACTGAC 849

Qy 1104 CTGGACACCCGCAACTGTACAGTACCTCTGTGTACACAGTCTCTGGCCCTGAGGAC 1163

Db 850 CTGGACACCCGCAACTGTACAGTACCTCTGTGTACACAGTCTCTGGCCCTGAGGAC 909

Qy 1164 GTGGCCCTCTATGTGGCCCTCTATGCGCCGTGGCCGTGCTGCTGTGCTGTGTC 1223

Db 910 GTGGCCCTCTATGTGGCCCTCTATGCGCCGTGGCCGTGCTGCTGTGCTGTGTC 969

Qy 1224 CTCATCTCGTTTATTTGCGGAGAGAGGGGCTGGACTCAGATGTGCTGACTGCTGC 1283

Db 970 CTCATCTCGTTTATTTGCGGAGAGAGGGGCTGGACTCAGATGTGCTGACTGCTGC 1029

Qy 1284 ATTCTCACCTCAGGCTTCCAGCCCGTTCAGATCAAGCCCGAGCAAGACCAACCCCAT 1343

Db 1030 ATTCTCACCTCAGGCTTCCAGCCCGTTCAGATCAAGCCCGAGCAAGACCAACCCCAT 1089

Qy 1344 CTGCTCACATCCAGCGGACCTCAG---CACACACACCTACACAGGGCAGTCTCTGT 1400

Db 1090 CTGCTCACATCCAGCGGACCTCAGCACACACACACACCTACACAGGGCAGTCTCTGT 1149

Qy 1401 CCCCAGGAGTGGGCGGCGGACACACTGACACACAGTCTCTCCACCTCTGAGGCCGAGGTT 1520

Db 1150 CCCCAGGAGTGGGCGGCGGACACACTGACACACAGTCTCTCCACCTCTGAGGCCGAGGTT 1269

Qy 1521 GTCTCCGCTCTCACACAGAACTACTTCGCTCTCCGCTCTCCGCGGACACACAGCAATG 1580

Db 1270 GTCTCCGCTCTCACACAGAACTACTTCGCTCTCCGCTCTCCGCGGACACACAGCAATG 1329

Qy 1581 ACCTATGGACCTTCAACTCTCTGGGCGCGGCTGATGATCCCTAATACAGGATACAG 1640

Db 1330 ACCTATGGACCTTCAACTCTCTGGGCGCGGCTGATGATCCCTAATACAGGATACAG 1389

Qy 1641 CTCCTCATCCCCAGATGCCATACCCGAGGAAAGATCTATGATCTTACCTCAGCTG 1700

Db 1390 CTCCTCATCCCCAGATGCCATACCCGAGGAAAGATCTATGATCTTACCTCAGCTG 1449

Qy 1701 CACAAGCCGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760

Db 1450 CACAAGCCGGAAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1509

Qy 1761 GTTAGCTGTGACCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820

Db 1510 GTTAGCTGTGACCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1569

Qy 1821 TGTGGGAGCCGACCCCTGACAGCTGGAGCTGCGCTCAAAAGAGTGTGCGAGGGC 1880

Db 1570 TGTGGGAGCCGACCCCTGACAGCTGGAGCTGCGCTCAAAAGAGTGTGCGAGGGC 1629

Qy 1881 AGCTGGAGCAGGATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940

Db 1630 AGCTGG---AGGATGTGTGACCTTGGCGGAGGAGGCGGCTGCTGCTGCTGCTGCTGCT 1686

Qy 1941 CAGCTGGAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000

Db 1687 CAGCTGGAGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746

Qy 2001 GGAGAGGCGCTCAGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2060

Db 1747 GGAGAGGCGCTCAGCGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1806

Qy 2061 GCCTGCACTCCCTCGAGTACAACTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2120

Db 1807 GCCTGCACTCCCTCGAGTACAACTCCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1866

Qy 2121 CTCACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2180

Db 1867 CTCACAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1926

Qy 2181 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2240

Db 1927 GTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1986

Qy 2241 TCCCTGTGGAAGTAAGTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2300

Db 1987 TCCCTGTGGAAGTAAGTCTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046

Qy 2301 AATGGACCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2360

Db 2047 AATGGACCCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106

Qy 2361 AGTGACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2420

Db 2107 AGTGACCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2166

Qy 2421 ATCAACTTCAACATCAACCAAGGACACAAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2480

Db 2167 ATCAACTTCAACATCAACCAAGGACACAAGGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226

Qy 2481 GCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2540

Db 2227 GCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2286

Qy 2541 AAGATAATTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2600

Db 2287 AAGATAATTTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2346

Qy 2601 CAGAAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2660

Db 2347 CAGAAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2406

Qy 2661 ATGATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2720

Db 2407 ATGATCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2466

Qy 2721 GCAGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2780

Db 2467 GCAGCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2526

Qy 2781 TGCTGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2840

Db 2527 TGCTGAGGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2586

Qy 2841 AGGCAGAGCCGCGAGCAGGGGCTTCCCGCACACCGGGGAGA 2881

Db 2587 AGGCAGAGCCGCGAGCAGGGGCTTCCCGCACACCGGGGAGA 2627

RESULT 5

MMU487852

LOCUS Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene). 3992 bp mRNA linear ROD 24-SEP-2002

DEFINITION Mus musculus mRNA for netrin receptor Unc5h1 (Unc5h1 gene).

ACCESSION AJ487852

VERSION AJ487852.1 GI:22035783

KEYWORDS netrin receptor Unc5h1; Unc5h1 gene.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1

Engelkamp, D.

Cloning of three mouse Unc5 genes and their expression patterns at mid-gestation

Mech. Dev. 118 (1-2), 191-197 (2002)

JOURNAL

MEDLINE 2239710

PUBMED 12351186

REFERENCE 2 (bases 1 to 3992)

AUTHORS Engelkamp,D.  
 TITLE Direct Submission  
 JOURNAL Submitted (15-MAY-2002) Neuroanatomy, Max Planck Institute for Brain Research, Deutscherordenstrasse 46, Frankfurt 60528, GERMANY  
 FEATURES  
 Location/Qualifiers  
 1..3992  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 1..3992  
 /gene="Unc5h1"  
 232..2928  
 /gene="Unc5h1"  
 /codon\_start=1  
 /product="netrin receptor Unc5h1"  
 /protein\_id="CAD32250.1"  
 /db\_xref="GI:22035784"  
 /db\_xref="GOA:O8K1S4"  
 /db\_xref="SPTREMBL:O8K1S4"  
 /translation="MAVRLGLPALGIVLTAWLRSGNAQSQSATVANPVPGANPDLLP  
 HFLVEPDDYIVNKEPVLLVCKAVPATQIFPKNGEWVQDVHVERSDGSSGLPTM  
 EVRINVSQKVEKFGLEEYWCQVAMSSGTTKQKAYIRIAYLRKNPEQRLAKEV  
 SLEQGITVPCRPPEGITPAEVEWLNEDLVPSLDPNVYITREHSLVQRQLADTAN  
 YTCVAKNIVARRSASAAVIVVNGSWTWSVCSAGRGWQKRSCTNPAPLN  
 CGAFCEGQNVQKTACATLCPVDGSMPSKWSACGLDCTHWSRECSDDPAPNGEEC  
 RGADLDTRNCTSDCLHTSSGPEDEVALYIGLVAVAVCLILLLLVLVLYCRKEGLDS  
 DVADSSILTSFGFQPVSIKESKADNPHELTIQDLSITTTTYYGSLCPRODGPSPFQL  
 SNHLLSPISGGRHTLHSSPTSEADVFSLRSTQNYFSLPRGTSNMAYGTFNFGG  
 RUMINTGTSLLIPDAIPDAIPRIYIYTHKPEDVRLPLAGCOTLLSPVSCGPPGV  
 LITRPFVILAMHCGSPSPDSWSLRKQSCGSEWEDVHLGSESPSHLYCQLEAGAC  
 VYFTEOLGRFALVPAVATKRLRLILLFAPVACTSLSEYNIYRVLYHDPHDALEVV  
 QLEKLGSLQIODEPRLHFKDSYHNLRLSIHDVPSLSKSLVSYQETPFYHINWGT  
 QYTLCTFLERNASTSDSLACKVWVWQVGGQSFNFNITKTDFRAEMALASEG  
 GVPALVGSFAFKIPILIRQKIITSLDPPCSRGADWRTLAQKULHDSHLSFFASKPSPT  
 AMILNWEARHPFNGLQALAAVAGLQPDAGLFTVSEAC"

ORIGIN

Query Match 80.8%; Score 2327.8; DB 10; Length 3992;  
 Best Local Similarity 89.3%; Pred. No. 0;  
 Matches 2579; Conservative 0; Mismatches 292; Indels 16; Gaps 6;

QY 1 AGCTGGGGCTCCGGGCTGAGCGCT-AAAGCCGCCCTCCCGCCGGGGGGCCCGCGCC 59  
 Db 149 AGTGGGGCTCTCGGCTGAGCGCTGACAGCTCTCTCCCGCCGGGGGCCCTAGTCC 208  
 QY 60 GGGCCCGCCGCTGCGCCCGCGCGCATGGCGTCCGGCCGGCGCTGTGGCCAGCGCTC 119  
 Db 209 AGCCCGCTCGT-----CCGCCCGCGCATGGCGTCCGGCCGGCGCTGTGGCCAGCGCTC 264  
 QY 120 CTGGGCATAGTCTCGCGCTGGCTCCGGCTCGGGTGGCGGCTCGGTCAGAGTGCCACCGTG 179  
 Db 265 CTGGGCATAGTCTCTACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 324  
 QY 180 GCCAACCCAGTGGCTGGTGCACACCGGAGCTGCTTCCCGACCTTCTCGTGGAGCCCGAG 239  
 Db 325 GCCAACCCAGTGGCTGGTGCACACCGGAGCTGCTTCCCGACCTTCTCGTGGAGCCCGAG 384  
 QY 240 GATGTGTACATGCTGAAGAACAAAGCCAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 299  
 Db 385 GAGCTGTACATTTGTCGAAGAACAAAGCCAGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 444  
 QY 300 CAGATCTTCTCAAGTGCACCGGGAGTGGTGGCGGAGTGGTGGCGGAGTGGTGGCGGAGTGG 359  
 Db 445 CAGATCTTCTCAAGTGCACCGGGAGTGGTGGCGGAGTGGTGGCGGAGTGGTGGCGGAGTGG 504  
 QY 360 AGCAGAGCGGAGTGGTGGAGCGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGG 419  
 Db 505 AGCAGTGCAGGAGCGAGGATTTGCCAACCAATGAGGTTCGGATCAAGGTATCAAGGAG 564  
 QY 420 CAGGTGCAGAGTGGTGGGCTGGAGGAATATCTGGTGGAGTGGTGGAGTGGTGGAGTGG 479  
 Db 565 CAGGTGCAGAAAGTGGTGGGCTGGAGGAGTGGTGGAGTGGTGGAGTGGTGGAGTGGAGTGG 624

QY 480 TCGGGCACCACCAAGAGTGCAGAGGCTACATCCGATAGCCAGATTTGGCAAGAACTTC 539  
 Db 625 TCAGGAACACCAAGAGTGCAGAGGCTACATCCGATTTGCCATTTTGGCAAGAACTTC 584  
 QY 540 GAGCAGGAGCGCTGGCCAAAGAGGTGTCCTTGGAGCGGGCATCGTGTGCTGCTGCTG 599  
 Db 685 GAGCAGGAGCGCTGGCCAAAGAGGTGTCCTTGGAGCGGGCATCGTGTGCTGCTGCTG 744  
 QY 600 CCACCGAGGAGCATCTCTCCAGCCAGGTGGAGTGGTCCGAAACGAGGACCTGTGTGAC 659  
 Db 745 CCCCAGGAGGAATCCCCCAGCTGAGGTGGAGTGGTCCGAAATGAGGACCTGTGTGAC 804  
 QY 660 CCGTCCCTGGACCCCAATGTATACATCACCGGAGCAGACGCTGTGTGTGCGACAGGCC 719  
 Db 805 CCTCTCTCGACCCCAATGTGTATACACGAGGAGCAGACCTAGTGTGTGCGGAGGCC 864  
 QY 720 CGCCTTGTGTGACACCGCCAACTACACCTGCTGGTGGCCAAACATCGTGTGCAAGTGC 779  
 Db 865 CGCCTGGCGGACACTGCCAACTACACCTGCTGGTGGCCAAACATCGTGTGCGGCTGCC 924  
 QY 780 AGCGCTCGCTGCTGTGTATCTGTACGTGAACGGTGGTGGTGGTGGTGGTGGTGGTGG 839  
 Db 925 AGCGCTCTGCGGCGCTGCTATTTATGTGAACGGTGGTGGTGGTGGTGGTGGTGGTGG 984  
 QY 840 TCCGCTCTGAGCGCGCAGCTGTGGCGCGCTGGCAGAAACGAGCGGAGCTGCACCAAC 899  
 Db 985 TCGCTGTGAGTGGCCAGCTGTGGCGCTGGTGGCAGAAACGAGCGGAGCTGCACCAAC 1044  
 QY 900 CCGGCGCTCTCAACGGGGCGCTTCTGTGTGAGGGGCAAGATGTCCATCAGCGCACCGTC 959  
 Db 1045 CCGGCACTCTCAACGGGGCGCTTCTGTGTGAGGGGCAAGATGTCCAGAAAAACAGCGCTGC 1104  
 QY 960 TCTCTCTGCTTGTCTGTGAGCGGAGCTGAGCGCGCTGGAGCGAGTGGTGGCGCTGT 1019  
 Db 1105 GGCATCTGT--TGCCCCAGTGTGAGTGGAGCTGGAGCCCAATGAGTAAAGTGGTGGTGG 1161  
 QY 1020 GGGCTGGAGTGCACCCACTGGCGGAGCGCTGAGTGTCTTGACCCAGCACCCCGCAACGGA 1079  
 Db 1162 GGGCTTGTACTGACCCACTGGCGGAGCGGGAGTGTCTCGAACCAGCGCCCGCAACGGA 1221  
 QY 1080 GGGGAGGAGTGCAGGGCACTGACCTGGACACCGCAACTGTACCAAGTCACTCTGTGTA 1139  
 Db 1222 GGTGAGGAGTGCAGGGGCTGCTGACCTGGACACCGCAACTGTACCAAGTCACTCTGTG 1281  
 QY 1140 GACAGTGTCTTGGCGCTGAGGAGCGCTGTATGTGGGCTCATCGCGCTGGCGCTG 1199  
 Db 1282 CACACCTCTTCCGGCGCCGAGGAGTGGTCTCTACATCGGCTCTGTCGCGTGGCGCTG 1341  
 QY 1200 TGCCTGGTCTGCTGCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1259  
 Db 1342 TGCCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1401  
 QY 1260 GACTCAGATGTGGTGAATCTGCTCAATCTCACTCAGGCTTCCAGCGCTGAGCATCAAG 1319  
 Db 1402 GACTCAGAGTGGTGAATCTGCTCAATCTCACTCAGGCTTCCAGCGCTGAGCATCAAG 1461  
 QY 1320 CCGCAGAAAGCAGAACCCCATCTGCTCACCATCCAGCGGAGCTCAGCACCCACCC--- 1376  
 Db 1462 CCGCAGAAAGCAGAACCCCATCTGCTCACCATCCAGCGGAGCTCAGCACCCACCCACG 1521  
 QY 1377 ACCACCTACAGGAGGAGTCTGTCTCCCGGAGGATGGGCGCCAGCCCGCCAGTTTCCAGCTC 1436  
 Db 1522 ACCACCTACAGGAGGAGTCTGTCTCCCGGAGGATGGGCGCCAGCCCGCCAGTTTCCAGCTC 1581  
 QY 1437 ACCAATGGGAGCTGTGTAGCGCCCTGGGTGGGCGCGGCACACATGACACCAAGCTCT 1496  
 Db 1582 TCTAATGGTCACTGTGTAGCGCCCTGGGTGGGCGCGGCATACGCTGACCAAGCTCT 1641  
 QY 1497 CCGACCTCTGAGCGGAGGATGCTCTCCGCTCTCCACCCAGAACTACTTCCGCTCC 1556  
 Db 1642 CCGACCTCTGAGGAGTGGAGTCTGCTCTCCGCTCTCCACCCAGAACTACTTCTGCTCT 1701  
 QY 1557 CTGCCCCGAGGAGCAGCAACATGACCTATGGGAGCTTCAACTTCTCGGGGCGCGGCTG 1616

Db 1702 CTGCCCCGCGGTACCAAGCAATGCGCTATGCGACCTTCAACTCTCTCGGGGCGCGCTG 1761  
QY 1617 ATGATCCCTTAATACAGGTATCAGCTCTCTATCCCTCCCAAGATGCATACCCGAGGGAAG 1676  
Db 1762 ATGATCCCTTAAACAGGAATCAGCTCTCTATACCCCGGAGCCATCCCGGAGAAAG 1821  
QY 1677 ATCTATGAGATCTACCTCAGCTGCAACAAGCGGAAGAGTGGGTTGCCCTAGCTGGC 1736  
Db 1822 ATCTACGAGATCTACTCTACTCTGCAACAAGCAGAGAGTGGGTTGCCCTAGCTGGC 1881  
QY 1737 TGTACAGCCCTGCTGAGTCCCAATCGTTAGCTGTGGACCCCTCGGCTCTGCTCACCCTG 1796  
Db 1882 TGTACAGCCCTGCTGAGTCCCAATCGTTAGCTGTGGGCCCCAGGAGTCTGCTCACCCTG 1941  
QY 1797 CCAGTCATCTGCTGATGACCACTGTGGGAGCCCGAGCCCTGACAGCTGGAGCTGGCC 1856  
Db 1942 CCAGTCATCTTGGCAATGACCACTGCGGGAGCCCGAGCCCGAGCTGGAGCTGGCC 2001  
QY 1857 CTCAAAAGCAGCTGCTGCGAGGCGAGCTGGGAGCAGATGTGCTGCACCTGGCGGAGGAG 1916  
Db 2002 CTCAAAAGCAGCTGCTGAGGGCAGCTGGG---AGGACGTGCTGCACCTTGTGAGGAG 2058  
QY 1917 GCGCCCTCCCACTCTACTACTGCGAGCTGGAGGCGAGTGCCTGTACTCTTCCACCGAG 1976  
Db 2059 TCGCCCTCTCATCTCTACTACTGCGAGCTGGAGGCGGCGCTGTATGTCTTCCACCGAG 2118  
QY 1977 CAGCTGGGCGGCTTTGCCCTGTGGAGAGGCGCCCTCAGCGTGGCTGCCAAGCGCCTC 2036  
Db 2119 CAGCTAGGCGGCTTTGCCCTGTGGAGAGGCGCCCTCAGCGTGGCTGCCAAGCGCCTC 2178  
QY 2037 AAGCTGCTTCTTTGCGCGGTGGCTGCACCTCCCTCGAGTACAACATCCGGGTCTAC 2096  
Db 2179 AGGCTCCTTCTGTTTGGCCCTGTGGCTGTAGTCCCTCGAGTACAACATCCGAGTCTAC 2238  
QY 2097 TGCCTGCATGACACCCACGATCGACTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2156  
Db 2239 TGCCTGCACGACACCCACGATCTCTCAAGGAGTGGTGCAGCTGGAGAGCAGCTGGGT 2298  
QY 2157 GGACAGCTGATCCAGAGGCCAGGCTCTGCACTTCAAGGACAGTTACCAACCTGGC 2216  
Db 2299 GGACAGCTGATCCAGAGGCCGCTGTCTGCACTTCAAGGACAGTTACCAACCTAGCT 2358  
QY 2217 CTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTCTGTGAGTACACAG 2276  
Db 2359 CTGTCCATCCAGAGTGCACAGCTCCCTGTGGAAGAGCAAGCTCTGTGAGTACACAG 2418  
QY 2277 GAGATCCCTTTTATCAATCTGGAATGSCAGGAGCGGTACTTGCACCTGCACCTTACC 2336  
Db 2419 GAGATCCCTTTTATCAATCTGGAATGSCAGTCACTCAGCAGTATCTGCACCTTACC 2478  
QY 2337 CTGGAGCGTGTGAGCCCGCAGACTAGTGCAGCTGGCGCTGCAGCTGGGTGGGAGGTG 2396  
Db 2479 CTGGAGCGCGTCAATGCGCAGCACAGCAGCTGGCGCTGCAAGGTGGGTGGGAGGTG 2538  
QY 2397 GAGGCGAGCGGCGAGCTTTCAGATCAACTTCAACATCAACAGGACACAGGTTTGTCT 2456  
Db 2539 GAGGAGATGAGCAGAGCTTCAATCACTTCACTTCACTAAGGACACAGGTTTGTCT 2598  
QY 2457 GAGCTGTGGCTCTGGAGAGTGAAGCGGGGTGCCAGCCCTGGTGGGCCCCAGTGCCTTC 2516  
Db 2599 GAAATGCTGGCTCTGGAGAGTGAAGGGGGGTGCCAGCCCTGGTGGGCCCCAGTGCCTTC 2658  
QY 2517 AAGATCCCTTCTCATTTCCGCGAGAGATATTTCCAGCTGGACCCACCTGTAGGCGG 2576  
Db 2659 AAGATCCCTTCTCATTTCCGCGAGAGATATTTCCAGCTGGACCCACCTGTAGGCGG 2718  
QY 2577 GGTGCCGAGCTGGCGACTCTGGCCCGAGAAATCCACCTGGACAGCACTCTCAGCTTCTTT 2636  
Db 2719 GCGCCGAGCTGGGAACTCTAGCCCGAGAACTTCACTTGGACAGCACTTCTAGCTTCTTT 2778  
QY 2637 GCCTCAAGCCCGAGCCCGACAGCCATGATCTCAACCTGTGGGAGCGCGGCACTTCCCC 2696

Db 2779 GCTTCAAGCCCGCCTACAGCATGATCTTCAACTATGGAGCGCGGCACTTCCCC 2838  
QY 2697 AACGGCAACTCAGCGAGCTGGCTGCGAGCAGTGGCTGGACTGGGCGCAGCCAGACGCTGGC 2756  
Db 2839 AACGGCAACTCAGCGAGCTGGCTGCGAGCAGTGGCTGGACTGGGCGCAGCCAGATGCTGGC 2898  
QY 2757 CTCTTCAAGTCTGGAGGCTGAGTCTGAGGCGCGCCAGGCGCG--ACACTTACTCT 2814  
Db 2899 CTCTTCAAGTCTGAGAGCCGAGTCTGAGACCAAGCCAGCCGCTTACCGCTTACTCT 2958  
QY 2815 CACCAGCTTTGGCACCACCAAGGACAGGACAGGCGGACAGGCGCTTCCCCACACC 2874  
Db 2959 CACCAGCTTTGGCACCCTGCCAGGACAGGACAGGCGGCGCTTCCCCACACC 3018  
QY 2875 GGGGAGA 2881  
Db 3019 CCGGGA 3025

RESULT 6  
BD057524  
LOCUS  
DEFINITION  
Netrin receptors.  
ACCESSION  
BD057524  
VERSION  
BD057524.1 GI:22603130  
KEYWORDS  
JP 2001505062-A/1.  
SOURCE  
synthetic construct  
ORGANISM  
artificial sequences.  
REFERENCE  
1 (bases 1 to 3014)  
AUTHORS  
Lavigne, M.T., Leonardo, D.E., Hinck, L., Masu, M. and Masu, K.K.  
TITLE  
Netrin receptors  
JOURNAL  
Patent: JP 2001505062-A 1 17-APR-2001;  
THE REGENTS OF THE UNIV OF CALIFORNIA  
COMMENT  
PN JP 2001505062-A/1  
PF 17-APR-2001  
PD 19-FEB-1998 JP 1998536840  
PR 19-FEB-1997 US 08/808982  
PI MARC TESSIER LAVIGNE, DAVID E LEONARDO, LINDSAY HINCK, MASAYUKI PI MASU,  
PI KAZUKO KEINO MASU  
PC C07K14/00, C07K14/00, C07K17/00, C07H21/02, C07H21/04, G01N33/53 CC  
Strandedness: Double;  
CC Topology: Linear;  
FH Key Location/Qualifiers.  
source  
1..3014  
/organism="synthetic construct"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:32630"

ORIGIN

Query Match 77.6%; Score 2235.8; DB 6; Length 3014;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

QY 87 ATGGCGCTCGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTGGCTC 146  
Db 1 ATGGCGCTCGGCGCGGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTGGCTC 60  
QY 147 CGCGCTCGGCTGGCCAGCAGAGTGCACCGTGGCCACCCAGTGCCTGGTGGCCACCG 206  
Db 61 CGTGGTTCGGGTGGCCAGCAGAGTGCACCGTGGCCAAATCCAGTGCCTGGTGGCCACCC 120  
QY 207 GACCTGCTTCCCCACTTCTCTGGTGGAGCCCGCAGGATGTGTACATCGTCAAGAAACAAGCA 266  
Db 121 GACCTGCTTCCCCACTTCTCTGGTGGAGCCCTGAGACGCTGTGTACATGTGTCAAGAACAGCG 180  
QY 267 GTGCTGCTGTGTGGAAGCGCGTGGCCCGCCACGAGATCTTTCTTCAAGTGAACGGGAG 326  
Db 181 GTGTTGTGTGTGCAAGGCTGTGCTGCCACCCAGATCTTCTTCAAGTGAATGGGAA 240  
QY 327 TGGGTGGCCAGGTGGACCACTGATCGAGGCGACACAGACGGGAGCAGTGTGTGAGCCG 386





QY	2544	ATAATTTCCAGCTGGAGCCACCTCTAGTGGGGGTGCCGACTCGCGGACTCTGGCCAG	2603
Db	2455	ATCATGCGCATCTGGAGCCCACTCTGACCGCGGGCGCGACTGGAGAACTCTAGCCCG	2514
QY	2604	AAACTCCACCTGGACAGCCATCTCAGCTTTCTTTTCCTCCAAAGCCCAAGCCCAAGCCCATG	2663
Db	2515	AAACTCCACCTGGACAGCCATCTTTAGCTTCTTTTCCTCCAAAGCCCAAGCCCATG	2574
QY	2664	ATCCTCAACCTGTGGAGGGCGGCACTTCCCCAAACGGCAACCTCAGCAGCTGGCTGCA	2723
Db	2575	ATCCTCAACCTGTGGAGGGCAAGCACTTCCCCAAACGGCAACCTCAGCAGCTGGCTGCA	2634
QY	2724	GCAGTGGCTGACTTGGGCGACAGAGCTGGCCCTCTTTCACAGTGTGGAGGCTGAGTGC	2783
Db	2635	GCTGTGGCGGACTTGGGCGAACCAGATGCTGGCCCTCTTTCACAGTGTGGAGGCTGAGTGC	2694
QY	2784	TGAGGCGGCGCAGGCGCG--ACACCTACACTCTCACCAGCTTTGGCACCCCAAGGACA	2841
Db	2695	TGAGACCAAGCCAGGCGCGTAATGCTACATCTCACCAGCTTTGACACCTGCCAGGACA	2754
QY	2842	GGCAGAGCCGACAGGCGCCCTTCCCCACACCGGGG	2878
Db	2755	GGCAAAACCAAGAGGCGCCCTTCCCCACACCGGGG	2791
RESULT 7			
LOCUS	AX268596	2697 bp	DNA linear PAT 29-OCT-2001
DEFINITION	Sequence 15 from Patent W00175440.		
ACCESSION	AX268596		
VERSION	AX268596.1	GI:16541710	
KEYWORDS			
SOURCE	Rattus sp.		
ORGANISM	Rattus sp.		
REFERENCE	1	Cochran,S.W., Paterson,G.Y., Ohashi,Y.W., Morris,B.Y. and Pratt,J.Y.	
AUTHORS		Schizophrenia related genes	
TITLE		Patent: WO 0175440-A 15 11-OCT-2001;	
JOURNAL		WELFIDE CORPORATION (JP)	
FEATURES		Location/Qualifiers	
source		1. 2697	
		/organism="Rattus sp."	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:10118"	
ORIGIN			
Query Match		76.0%;	Score 2189.4; DB 6; Length 2697;
Best Local Similarity		88.9%;	Pred. No. 0;
Matches 2403; Conservative		0;	Mismatches 291; Indels 9; Gaps 3;
QY	87	ATGCGCGTTCGGCGCGGCTGTGGCAGCGCTCTCTGGGCATGATCTCTGGCGTTGGCTC	146
Db	1	ATGCGCGTTCGGCGCGGCTGTGGCAGTCTCTGGGCATGATCTCTGGCGTTGGCTC	60
QY	147	CGCGGCTTCGGGTGCCAGAGAGTGGCACCGTGGCCAAACCCAGTGTGGTGCACACCG	206
Db	61	CGTGTTCGGTTCGCCAGAGAGTGGCACCGTGGCCAAACCCAGTGTGGTGCACACCG	120
QY	207	GACCTGTTCCCACTTCTGTGGGCGCGAGATGTGTACATCGTCAAGAACAGGACA	266
Db	121	GACCTGTCGCCACTTCTGTGTAGAGCTGTAGAGCTGTACATTTGTCAAGAACAGGCG	180
QY	267	GTGCTGCTGTGTGAAGGCGCGTGGCCGCAACCGAGATCTTCTCAAGTGCACCGGGAG	326
Db	181	GTGTGTGTGTGTGCAAGGCTGTGGCTGCCACCCAGATCTTCTTCAAGTGCATGGGAA	240
QY	327	TGGGTGCGCGAGTGGACCAAGTGTGACGCGCAGCAGACGGAGAGTGGTGAAGCG	386



```
1464 GGTGGCGCCGCCACACACTGACACACAGCTCTCCCACTCTGAGCGCGAGGAGTTGCTC 1523
1378 GGGAGTGGCGGCATACGTTGACACACAGCTCACCACTCTGAGGCTGAGACTTCGTC 1437
1524 TCCCGCTCTCCACACAGAACTACTTCGCTCCCTGCCCCGAGGACACAGCAACATGACC 1583
1438 TCCCGCTCTCCACCAAACTACTTCGTTCCCTGCCCCGAGGACACAGCAACATGACC 1497
1584 TATGGACCTTCAACTTCTCGGGGCGGGCTGATGATCCCTTAATACAGGTATACGCTTC 1643
1498 TACGGGACCTTCAACTTCTCGGGGCGGGCTGATGATCCCTTAATACAGGGATCAGCTTC 1557
1644 CTATATCCCCAGATGCCATACCCGAGGAGATCTATGAGATCTACCTGACGTGCAC 1703
1558 CTATATCCCCCGGATGCCATACCCGAGGAGATCTACGAGATCTACCTGACATGAC 1617
1704 AAGCCGGAAGAGCTGAGGTTGGCCCTAGCTGCTGTACAGCCTGTGAGTCCCATCGTT 1763
1618 AAGCCAGAGAGCTGAGGTTGGCCCTAGCTGCTGTACAGCCTGTGAGTCCCATCGTT 1677
1764 AGCTGTGACCCCTGGGCTGCTGTACCCGCGCAGTCACTCTGCTATGACCACTGT 1823
1678 AGCTGTGGCCGCCAGGAGTCTGCTACCCGCGCAGTCACTCTGCAATGGACCACTGT 1737
1824 GGGAGCCCGCCTGACAGCTGGAGCTCGCCTCAAAAGCAGTGTGCGAGGCGACG 1883
1738 GGAGAGCCCGCCTGACAGCTGGAGTCTGCGCCTCAAAAGCAGTGTGCGAGGCGACG 1797
1884 TGGGAGCAGAGTGTGTCACCTGGCGGAGGAGCGCCCTCCCACTCTACTACTGCCAG 1943
1798 TGGG---AGGATGCTGACCTTGTGAGGAGTCACTTCCCACTCTACTACTGCCAG 1854
1944 CTGGAGCCAGTGCCTGTACGTTTACCGAGCAGCTGGCGCGCTTGCCTGTGGGA 2003
1855 CTGGAGCCCGCCTGCTATGTTTACCGAGCAGCTGGCGCGCTTGCCTGTGGAGGA 1914
2004 GAGGCCCTCAGGTGCTCGCGCAAGCGCTCAAGCTGCTTCTGTTGCGCGGTGGC 2063
1915 GAGGCCCTCAGGTGCTCGCCACCAAGCGCTCAGGCTCTCTGTTGCTCCGTTGGC 1974
2064 TGCACTCCCTCGAGTACCAATCCCGGTCTACTGCTGATGACACCCACGATGACATC 2123
1975 TGTAGTCTCTTGTAGTACCAATCCGAGTACTGCTACACGACCCACGAGCTCTC 2034
2124 AAGAGGTGGTGTGAGAGAGAGCTGGGGGACAGCTGATCAGAGGACACCGGCTC 2183
2035 AAGAGGTGGTGTGAGAGAGAGCTAGGTGAGCAGCTGATCAGGAGGCTCGCTC 2094
2184 CTGCACTTCAAGGACAGTTACCAACCTGCGCTATCCATCCAGATGTGCCAGCTCC 2243
2095 CTGCACTTCAAGGACAGTTACCAACCTGAGTCTCTCATCCAGGAGTCCAGCTCC 2154
2244 CTGTGAAGAGTAAGCTCTCTGTGAGTACAGGAGATCCCTTTTATCAGATCTGGAAT 2303
2155 CTGTGAAGAGTAAGCTCTGTGAGTACAGGAGATCCCTTTTATCAGATCTGGAAT 2214
2304 GGCACGCGGTGACTTGTGACTGCTTACCTTGGAGGCTGTGAGCCCGACACTAGT 2363
2215 GGCACCGCAGTATCTGCACTGCTTACCTTGGAGGCTGTGAGCCCGACACTAGT 2274
2364 GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGCGGCGAGAGCTTTCAGATC 2423
2275 GACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGGAGGAGATGGGAGAGCTTTCAGATC 2334
2424 AACTTCAACATCAACAGGACAAAGTGTGAGTGTGCTGTGAGAGTGAAGCG 2483
2335 AACTTCAACATCAACAGGACAAAGTGTGCTGATTTGCTGCTTGGAGAGTGAAGG 2394
2484 GGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCGAAG 2543
2395 GGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCGAAG 2454
```

```
2544 ATATTTCCAGCTGGACCCACCTGTAGGGGGGTGGCGACTGGCGGACTCTGGCCAG 2603
2455 ATCATCGCAGTCTGGAACCCCTCTGAGCGGGCGCGACTGGAGACTCTAGCCAG 2514
2604 AAATCCACCTGGACAGCACTCTAGCTTCTTGTCTCAAGCCAGCCCAAGCCATG 2663
2515 AAATCCACCTGGACAGCACTCTAGCTTCTTGTCTCAAGCCAGCCCTACAGCAATG 2574
2664 ATCTCAACCTGTGGAGGGCGGCACTTCCCAAGCGCAACCTCAGCCAGCTGGCTGA 2723
2575 ATCTCAACCTGTGGAGGGCGGCACTTCCCAAGCGCAACCTCAGCCAGCTGGCAGA 2634
2724 CGAGTGGCTGGACTGGCGCAGCAGCAGCTGGCCCTTTCACAGTGTGCGAGGCTGAGTGC 2783
2635 GTGTGGCGGAGTGGCCCAACCATGCTGCGCTTTCACAGTGTGCGAGGCGGAGTGT 2694
2784 TGA 2786
2695 TGA 2697
```

RESULT 8  
RN087305  
LOCUS  
DEFINITION  
Rattus norvegicus transmembrane receptor Unc5H1 mRNA, linear ROD 15-MAY-1997  
ACCESSION  
U87305  
VERSION  
U87305.1 GI:2055391  
KEYWORDS  
Rattus norvegicus (Norway rat)  
SOURCE  
Rattus norvegicus  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus  
REFERENCE  
1 (bases 1 to 2697)  
Leonardo,E.D., Hinc,L., Masu,M., Keino-Masu,K., Ackerman,S.L. and  
Tessier-Lavigne,M.  
Vertebrate homologues of C. elegans UNC-5 are candidate netrin  
receptors  
NATURE 386 (6627), 833-838 (1997)  
9721897  
9126742  
2 (bases 1 to 2697)  
Leonardo,E.D., Hinc,L., Masu,M., Keino-Masu,K. and  
Tessier-Lavigne,M.  
Direct Submission  
Submitted (28-JAN-1997) Anatomy, UCSF, 513 Parnassus, San  
Francisco, CA 94143-0452, USA  
LOCATION/Qualifiers  
1..2697  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10116"  
/tissue\_type="brain and ventral spinal cord"  
/dev\_stage="18 day embryo and 13 day embryo"  
1..2697  
/codon\_start=1  
/product="transmembrane receptor UNC5H1"  
/protein\_id="AA057678.1"  
/db\_xref="GI:2055392"

CDS  
/translation="MAVRPGLMPVLIGIVLAALWLGSGAQQSATVANVPGANPDLLP  
HFLPEPDIYIVNKPVLVKAIPATQIFFKNGEWMVQVDHVIERTSDSSGUPM  
EYRINVRQVKEVQVWSSGTTKQKAYIRIAYLRNFEQELAKEV  
SLEQGLVFCRPEGIPPAEVEWLNEDLVDPNVIYTHESLVVRQARLAFAN  
YTCVANIVARRSTAAIVVVGSGMTWTSVCSACGRWQKRSCTNPAPLN  
GAFCEGVNQKTACATLCPVGSWSMSKWSACGLDCTHWSRECSDPAPRNGREC  
RGADLNRCTSDLCUHTASCPEDVALYGLVAVACLFLLLALGLIYCRKKEGLDS  
DVADSLTSGFPVSIKPSADNPHLLTIQPDLSITTTTYSGLSCRGQSPKFLQ  
SNHGLSLPGSGRHTLHSSPTSEADFSRLSTQNYFRSLPRGTGNMAYTNFNLGG  
RLMIPNTGSLIIPDAIPRGKIYIYTLHKPEDVRLPLAGCOTLLSPVYSCGPGV  
LITRPVILAMDHCGEPPDSWLSRLKKQCEGSEWDLHLGEBSPSHLYYCOLAEAGAC  
VYPTGOLGRPALVGEALSVAATKRLFLFAPVACTSLEVINIRVYCLHPTDHALKEV  
QEKQGGGLIQRPVLFHFDKYNLRLSIHDPVSSILWKSLLVYQEIIPPYHNGT  
QQLYHCTFLERINASTDLACKVWVWVQVEGQGFNFINFTKTRFALLALESEG

ORIGIN	Query Match Best Local Similarity 76.0%; Score 2189.4; DB 10; Length 2697; Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;	GVPALVGSFAFKIPPLINQKIIASLDPPCSRGADWRITLQAKLHLDLSHLSFFASKPSPT AMILNLWEARHPNGLGOLAAAVAGLQPDAGLFTVSEAE"
QY	87 ATGGCCGTCGCGCCCGGCGCTGTGGCCAGCGGTCTCTGGGCATAGTCTCGCGCTGGGCTC 146	958 CGCGAGTGTCTGTACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG 1017
DB	1 ATGGCCGTCGCGCCCGGCGCTGTGGCCAGCGGTCTCTGGGCATAGTCTCGCGCTGGGCTT 60	1107 GACACCCGCAACTGTATACAGTGAACCTCTGTGTATACAGTGTCTTCTGGCCCTTGGAGACGTG 1166
QY	147 CGCGCTCGGCTGCCACGACAGTGCACCGTGGCCAAACCCAGTGCCTTGTGCCAACCCG 206	1018 GACACCCGCAACTGTATACAGTGAACCTCTGTGTATACAGTGTCTTCTGGCCCGGAGACGTG 1077
DB	61 CGTGGTTCGCGTGGCCACGAGTGCACCGTGGCCAAATCCAGTGCCTGGCCGCGTGCACCC 120	1167 GCCCTCTATGTGGSCCTCATCGCGTGGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1226
QY	207 GACCTGTTCCTCCACTTCTCGTGAGCCGCGAGGATGTATCATCTGTCAAGAAACAGGCA 266	1078 GCTCTCTACATCGCGCTTGTGCTGTGCTGTGCTGTCTTCTGCTGTGCTGTGCTGTGCTGCT 1137
DB	121 GACCTGTTCCTCCACTTCTCGTGAGCCTGAGGACGTGTATCTTCAAGAAACAGGCG 180	1227 ATCTCTGTTTATTCGCGAAAGAGGAGGCGCTGAGCTCAGATGTGGGTGCTGACTCTCTCCATT 1286
QY	267 GTGCTGTGTGTGCAAGGCGGTGCCGCCACGAGATCTTCTTCAAGTGCAAACGGGAG 326	1138 GGACTCATTTACTGTCTCAAGAGAGAGGCGTGGACTTCGATGTGGCGGACTGTGTCTCATC 1197
DB	181 GTGTGTGTGTGTGCAAGGCTGTGCTCTGCCACCCAGATCTTCTTCAAGTGCAAATGGGAA 240	1287 CTACCTCAGGCTTCAGCCCGTTCAGCATCAAGCCCAAGCAAAAGAGCAAAACCCCATCTG 1346
QY	327 TGGGTGCCCAGGTGGACCATGTGATCGAGCGCAGACAGACGGGAGCAGTGTGAGCG 386	1198 CTCACCTCGGCGCTTCAGCGCTGTGAGCATCAAGCCCAAGCAAAAGAGCAAAACCCCATCTG 1257
DB	241 TGGGTGCCCAGGTGCATACGTAATTAAGCGCAGACCCAGATCTTCTTCAAGTGCAAATGGGAA 300	1347 CTACCATTCAGCGGACCTCAGCACACC--ACCACTTACAGGCGAGTCTCTGTGCC 1403
QY	387 ACCATGGAGTCCGATTAAATGTCTCAAGGACAGAGTCGAGAGGTGTTCGGGCTGGAG 446	1258 CTACCATTCAGCGGACCTCAGCACCACTATCACCTTACAGGCGAGTCTCTATGTTG 1317
DB	301 ACCATGGAGTCCGATTCAACGTATCGAGGACAGAGGTAGAGAAAGTGTTCGGGCTGGAG 360	1404 CGCAGGATGGGCCAGCCCCAAGTTCAGCTCACCATGGGCACTGTCTCAGCCCCCTG 1463
QY	447 GAATACTGTGCCAGTGGCATGGAGTCTCTCGGSCACCAACAAAGAGTCAGAAAGCC 506	1318 AGGCAGGATGACCCAGCCCCAAGTTCACGCTCTCTAATGCTCAGCTCTCAGCCACTG 1377
DB	361 GAATACTGTGCCAGTGTGGCATGGAGTCTCTCGGSCACCAACAAAGTTCAGAAAGCC 420	1464 GGTGGGCGCCGACACACATGACACAGCTCTCCACCTCTGAGGCCGAGGAGTTCGTC 1523
QY	507 TACATCCGATAGCCAGATTCGCAAGAACTTCGAGCAGGAGCGCTGGCCAAAGAGTG 566	1378 GGGAGTGGCGCCATACGTTTGACACAGCTCACCACCTCTGAGGCTGAGGACTTCGTC 1437
DB	421 TACATCCGATAGCTTATTTGCGCAAGAACTTTCAGCAGGAGCGACTGGCCCAAGAGTG 480	1524 TCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCCGAGGACCAAGCAATGAC 1583
QY	567 TCCCTGAGCAGGACATCGTCTGCCCTCGCTCCACCGAGGACATCCCTCCAGCCGAG 626	1438 TCCCGCTCTCCACCCAAACTACTTTCGTTCCCTGCCCCGCGCACCAAGCAATGAGCC 1497
DB	481 TCATCGAGCAGAGCATTGTACTACTTGTGCGCCCGCAGAGGAATCCCGCCAGCTGAG 540	1584 TATGGGACCTTCAACTTCTTCGGGGCGCGCTGATGATCCCTAATACAGGTATCAGGCTC 1643
QY	627 GTGAGTGGCTCCGAAACGAGGACCTGTGTGACCCCGTCCCTGGACCCCAATGTATCATC 686	1498 TAGCGGACCTTCACTTCTTCGGGGCGGCTGATGATCCCTAATACGGGATCAGGCTC 1557
DB	541 GTGAGTGGCTTCGAAATAGGACCTCTGTGACCCCTCCCTCGATCCCAATGTGTATC 600	1644 CTATCCCCCGAGATGCGATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703
QY	687 ACSCGGAGACACGCTGTGTGCGACAGCCCGCTTGTGACACGGCCAACTACAC 746	1558 CTATATCCCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCAGCTGCAC 1617
DB	601 ACGGGGAGACACGCTAGTGTGCGTCAAGCCCGCTGCGCGACACCGGCCAACTACAC 660	1704 AAGCCGGAAGAGTGGAGTTGCCCTAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1763
QY	747 TCGTGGCCCAAGAACATCGTGGACGTCGCGGAGCGCTCCGCTGTGTGTGTGTGTGTGT 806	1618 AAGCCAGAGAGTGGAGTTGCCCTAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1677
DB	661 TGTGTGCCCAAGAACATCGTAGCCCTGCGCGAGCCTCTGAGCGGCTATGTTTAT 720	1764 AGCTGTGGACCCCTTCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1823
QY	807 GTGAACCGGTGGTGTGACGTGACCGAGTGTGTGCTGTGACGCGCAGCTGTGGCGC 866	1678 AGCTGTGGGCCCCCAGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1737
DB	721 GTGAACCGGTGGTGTGACGTGACCGAGTGTGTGCTGTGACGCGCAGCTGTGGCGT 780	1824 GGGAGCCCGAGCCCTGACAGCTGGAGCTTGCGCTCAAAAAGAGTGTGCGAGGCGAG 1883
QY	867 GGTGGCAGAAACGGAGCCGAGCTGCACCAACCCGCGCTCTCAACGGGGCGCTTTC 926	1738 GGAGAGCCCGAGCCCTGACAGCTGGAGTCTGCGCTCAAAAAGAGTGTGCGAGGCGAGT 1797
DB	781 GGTGGCAGAAACGGAGCCGAGCTGCACCAACCCGCGCTCTCAACGGGGCGCTTTC 840	1884 TGGAGCAGAGTGTGCTGTGCTGTGCGAGGAGGCGCCCTCCCACTCTACTACTGCGCAG 1943
QY	927 TGTGAGGGCAGAAATGCCATGACCGGACCGTCTCTCTCTGTGTGTGTGTGTGTGTGT 986	1798 TGGG---AGGATGTGTGTGCTGTGCTGTGAGGAGTCACTTTCACCTCTACTACTGCCAG 1854
DB	841 TGTGAGGGCAGAAATGCCATGACCGGACCGTGTGGCCACTCTG---TGCCAGTGGATGG 897	1944 CTGGAGCCAGTCCCTGTCTTACGTTACCGAGAGCTGGGCGGCTTTCCTCTGTGGGGA 2003
QY	987 AGCTGAGCCCGTGGAGCAAGTGGTGGCTGTGGCTGTGAGTGCACCCACTGCGCGAGC 1046	1855 CTGGAGCCCGGCGCTGTCTATGCTTTCAGGAGAGCTGGGCGGCTTTCCTCTGTGAGGA 1914
DB	898 AGCTGAGTTCGTGGAGTAAAGTGTGAGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 957	2004 GAGGCCCTCAGGTGGCTGCCGCCAAGCGCTCTCAAGTGTCTTCTGTGTGCGCGGTGGCC 2063
QY	1047 CGTGTGTGTCTGACCCAGCACCCCGCAACCGGAGGAGAGTGCACGGGCACTGACCTG 1106	1915 GAGGCCCTCAGGTGGCTGCCACCAAGCGCTCAGGCTCCTTCTGTGTGCTCCCGTGGCC 1974
		2064 TGCACTTCCCTCGAGTACAACTATCGGGTCTACTGCTGTGATGACACCCCAAGTCACTC 2123
		1975 TGTGTGTCTTGTGAGTACAACTCGGAGTGTACTGCTTACAGCACCCACGACGCTCTC 2034
		2124 AAGGAGGTGGTGTGAGAGAGAGCTGGGGGACAGCTGATCCAGGAGCAGCGGTC 2183

Db 2035 AAGGAGGTGGTGCAGCTGGAGAGCAGCTAGTGTGACAGCTGATCCAGGAGCCTCGCGTC 2094

Qy 2184 CTGCACCTTCAAGGACAGTTACACAACTTGGCCCTTATCATCCAGATGTGCCAGCTCC 2243

Db 2095 CTGCACCTTCAAGGACAGTTACACAACTTGGCCCTTATCATCCAGATGTGCCAGCTCC 2154

Qy 2244 CTGTGGAGAGTAAGCTCTTGTGACGTACAGGAGATCCCTTTTATCATCTTGGAT 2303

Db 2155 CTGTGGAGAGTAAGCTCTTGTGACGTACAGGAGATCCCTTTTATCATCTTGGAT 2214

Qy 2304 GCACGCGAGCGTACTTGCACCTGACCTTACCTTGGAGCGTGTGAGCCGACCTAGT 2363

Db 2215 GCACGCGAGCGTACTTGCACCTGACCTTACCTTGGAGCGTGTGAGCCGACCTAGT 2274

Qy 2364 GACCTGGGCTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGAGCGGAGAGCTTTCAGATC 2423

Db 2275 GACCTGGGCTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGAGCGGAGAGCTTTCAGATC 2334

Qy 2424 AACTTCAACATCAACCAAGGACACAGGTTTGTGAGCTGCTGCTGTGGAGAGTGAAGG 2483

Db 2335 AACTTCAACATCAACCAAGGACACAGGTTTGTGAGCTGCTGCTGTGGAGAGTGAAGG 2394

Qy 2484 GGGGTCCCAAGCCTGCTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAAG 2543

Db 2395 GGGGTCCCAAGCCTGCTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAAG 2454

Qy 2544 ATAATTCAGCTGACCCAGCCCTGTAGCGGGGTGGCGACTGCGGAGCTCTGGCCAG 2603

Db 2455 ATCATCGCCAGTCTGGACCCAGCCCTGAGCGGGGCGCGCACTGGGAGAACTCTAGCCAG 2514

Qy 2604 AACTTCAACATCAACCAAGGACACAGGTTTGTGAGCTGCTGCTGTGGAGAGTGAAGG 2663

Db 2515 AACTTCAACATCAACCAAGGACACAGGTTTGTGAGCTGCTGCTGTGGAGAGTGAAGG 2574

Qy 2664 ATCTTCAACATCAACCAAGGACACAGGTTTGTGAGCTGCTGCTGTGGAGAGTGAAGG 2723

Db 2575 ATCTTCAACATCAACCAAGGACACAGGTTTGTGAGCTGCTGCTGTGGAGAGTGAAGG 2634

Qy 2724 CGAGTGGCTGAGTGGCGGCGAGCGAGCGTGGCTTTCAGATGTGGAGGCTGAGTGC 2783

Db 2635 GCTGTGGCGGAGTGGCGGCGAGCGAGCGTGGCTTTCAGATGTGGAGGCTGAGTGC 2694

Qy 2784 TGA 2786

Db 2695 TGA 2697

RESULT 9

BC058084

LOCUS BC058084

DEFINITION Mus musculus unc-5 homolog A (C. elegans), mRNA (cdna clone

BC058084

VERSION MG:66671 IMAGE:6813463), complete cds.

KEYWORDS

SOURCE MG.

ORGANISM Mus musculus (house mouse)

REFERENCE 1 (bases 1 to 3844)

AUTHORS Strausberg R.L., Feingold, E.A., Grouse, L.H., DeJong, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buettow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prance, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,

Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalish, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 3844)

Strausberg, R.

Direct Submission

Submitted (08-SEP-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)

Tissue Procurement: Dr. Jim Lin, University of Iowa

CDNA Library Preparation: M. Bento Soares, University of Iowa

CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hailao, Martin Krzyzanski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>

Series: IRAC Plate: 126 Row: b Column: 11

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 23346570.

Location/Qualifiers

1. 3844

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6"

/db\_xref="taxon:10090"

/clone="MG:66671 IMAGE:6813463"

/tissue\_type="Brain, mouse, 13.5, 14.5, 16.5, 17.5 dpc"

/clone\_lib="NIH BMAP\_FW0"

/lab\_host="DH10B"

/note="Vector: pYX-ASC"

1. 3844

/gene="Unc5a"

/db\_xref="LocusID:107448"

/db\_xref="MGI:894682"

252. 2780

/codon\_start=1

/product="Unc5a protein"

/protein\_id="AAH5084.1"

/db\_xref="GI:34784159"

/db\_xref="LocusID:107448"

/translation="MAVRPLMPALGIVLTAWLRGSGAQSATVANPVGANPDLLP

HLFVPEVIVKVKPVLLVCAPATQIFKNGEVRQVDHIVRSTGSGSLPTM

EVRIIVSRQVKVFCLEWVCOVWSSGTTKQAVIRIAYLRKNEQELAKEV

SLEQGLVLPFRPEGIPPAEVLVMDVLDPSLDNVNIVREHSLVVRQARLADTN

YTCVAKNIVARRSAGAAVIVVVGSGSPWKSACGLDTHWRECSDPARNGE

ECRGADLDTRNCTSDLCITSGSPDNVILVAVACLLILLVLLVLYCKRKEGL

DSVADSSILTSQFQVSIKPSKADNPILLITQDLSITTTTQGSICPRQDPSKPF

QLSNGHLLPSGLIGRTHLHSSPTSEAEFVSRSLSTQNVPRSLPRGTNNMAYGTNFI

GGRLMIPNTGISLIPDAPRGKIYEIVTLHKPDLVPLGCTLLPSIVSCGPP

GULLTPVILMDHCCPEPSDLSRLKQSCGSMEDVILGCEPSHLYVCOLEAG

ACVFTQELGRFALVGEALSVATKRLRLLPAPVACTSLEYNIRVYCLDTHDLKE

VVQLEKQLGQLIQEIPRHLFHKDSYHNLRLSHDVPSSLWKSLLVYSIQEIPFTHIWN

FEATURES

Source

gene

CDS

```
GTQYLHCTFTLLERNVASTDLACKVWVWQVGDGOSFNINFTKDTFPAEMIALES
EGVPAVGPSAFKIPFLIRQKILTSIDPCSRGADWRTLAOKHLDSLSPSPKPS
PTAMINLWEARHPNGLNQLAAAVAGLQDPDAGLFTVSEAC"
714. .926
/notes="IG; Region: Immunoglobulin"
/db_xref="CDD:smart00409"
misc_feature 984. .1130
/notes="TSPI; Region: Thrombospondin type 1 repeats"
/db_xref="CDD:smart00209"
misc_feature 1566. .1877
/notes="ZU5; Region: Domain present in ZO-1 and Unc5-like
necrin receptors"
/db_xref="CDD:smart00218"
misc_feature 2496. .2684
/notes="DEATH; Region: DEATH domain, found in proteins
involved in cell death (apoptosis). Alpha-helical domain
present in a variety of proteins with apoptotic functions.
Some (but not all) of these domains form homotypic and
heterotypic dimers"
/db_xref="CDD:smart00005"
ORIGIN
Query Match 70.1%; Score 2020.6; DB 10; Length 3844;
Best Local Similarity 84.1%; Pred. No. 0;
Matches 2429; Conservative 0; Mismatches 274; Indels 184; Gaps 6;
QY 1 AGCTGGGGCTCGGGCTGAGGGCT-AAAGCCGCTCCCGCGCGGGGGCCCGGCC 59
Db 169 AGGTGGGGCTCTCGGCTGAGGGCTGACAGCTCTCTCCCGCGCGGGGGCCCTAGTCC 228
QY 60 GCGCGCGCGCTGCGCGCGCGCCATGGCCGTCGCGCGCGCGCTGTGGCCAGCGTCC 119
Db 229 AGCCCGCTCGT---CCGCGCGCGCGCCATGGCGCTCGCGCGCGCTGTGGCCAGCGTCC 284
QY 120 CTGGSCATAGTCTCGCGCTTGGTTCGCGGCTCGGGTGCCCGAGAGTGCCACCGTG 179
Db 285 CTGGSCATAGTCTCTCACTGCGCTGGTTCGTGGTTCGGGTGCGCCAGAGTGCCACAGTG 344
QY 180 GCCAACCCAGTCTGTGTGCAACCCGACCTGCTTCCCACTTCTGTTGGAGCGCCGAG 239
Db 345 GCCAACCCAGTCTGTGTGCAACCCGACCTGCTTCCCACTTCTGTTGGAGCGCCGAG 404
QY 240 GATGTGTCATGTGTCAGAAACAGCCAGTGTGCTTGTGTGCAAGCGCGTGCCTCCACG 299
Db 405 GACGTGTACATTTGTCAAGAAACAGCCGCTGCTGTGTGTGCAAGCGTGTGCCTCCACG 464
QY 300 CAGATCTTCTCAAGTCAACGGGGAGTGGTGCGCGCAGGTGGACCACTGATCGAGCGC 359
Db 465 CAGATCTTCTCAAGTCAACGGGGAGTGGGTTCGCGCAGGTGCGATCACTGATGAAGCC 524
QY 360 AGCACAGCGGAGCAGTGTGTGAGCGGACCATGGAGGTCCGCATTAAATGCTCAAGGAG 419
Db 525 AGCACTGACGCGCAGCGGATTGCCAACCATGGAGGTCCGGATCAAGTATCAAGGCGAG 584
QY 420 CAGTCCAGAAAGTGTTCGGCTCGAGGAATPACTGGTCCAGTGTGCTGGCATGGAGTCC 479
Db 585 CAGTCCAGAAAGTGTTCGGCTCGAGGAGTACTGGTCCAGTGTGTGGCATGGAGTCC 644
QY 480 TCGGGCCACCCAGAGTCAAGAGCCCTACATCCGATAGCCAGATTGCGCAAGAACTTC 539
Db 645 TCAGGAACCAACAAAGCCAGAGGCTTACATCCGGATTGCTATTTGCGCAAGAACTTT 704
QY 540 GAGCAGGAGCGCTGGCCAGAGAGTGTCCCTGGAGCAGGCGCATCGTGTGCCCTGCGGT 599
Db 705 GAGCAGGAGCGCTGGCCAGAGAGTGTCACTGGAGCAAGGCATTGTGCTACCTGTGCG 764
QY 600 CCACCGAGGCGCATCTCCAGCCGAGTGGAGTGGCTTCGGAAACGAGGACTGGTGGAC 659
Db 765 CCCCAGGAAGGAATCCCCCAGCTGAGGTGAGTGGCTCCGAAATGAGGACCTCGTGGAC 824
QY 660 CCGTCCCTGACCCCAATGTATACATCAGCGGGAGCAGACCTGTGGTGGCAGAGGCC 719
Db 825 CCCTCCCTCGACCCCAATGTATACATCAGCGGGAGCAGACCTGTGGTGGCAGAGGCC 884
```

```
QY 720 CGCTTGTGTGACACGCGCAACTACACTGCGTGGCCCAAGAACATCGTGGCACCTCGCGC 779
Db 885 CGCTTGTGTGACACGCGCAACTACACTGCGTGGCCCAAGAACATCGTGGCACCTCGCGC 944
QY 780 AGCGCTTGTGTGACACGCGCAACTACACTGCGTGGCCCAAGAACATCGTGGCACCTCGCGC 839
Db 945 AGCGCTTGTGTGACACGCGCAACTACACTGCGTGGCCCAAGAACATCGTGGCACCTCGCGC 963
QY 840 TCGTGTGTCAGCGCGCAACTGCGTGGCGCGCTGCGAGAAACGGAGCGGAGTGCACCAAC 899
Db 964 ----- 963
QY 900 CCGGCGCTCTCAACGCGGGGCGCTTCTGTGAGGGGAGAGATGTCATGACCGCACCGTCC 959
Db 964 ----- 963
QY 960 TCCTCTCTGCTGTCTCTGTGACGCGCAGCTGGAGCCGCTGGAGCAAGTGGTGGCGCTGT 1019
Db 964 -----TTGTTTATGTGGATGGAGCTGGAGCCCATGAGTAAGTGGTTCAGCGCTGC 1013
QY 1020 GGGCTGGAGTGCACCCACTGCGCGGAGCGCTGAGTGTCTGTGACCCAGCACCCCGCAACGGA 1079
Db 1014 GGGCTGGAGTGCACCCACTGCGCGGAGCGCGGAGTGTCTGAGCCAGCGCGCCCGCAACGGA 1073
QY 1080 GGGGAGGAGTGCACGCGCAGCTGACCTGGACACCCCGCAACTGTACCACTGTCTGTGTA 1139
Db 1074 GGTGAGGAGTGCAGCGGCTGTGACCTGGACACCCCGCAACTGTACCACTGTCTGTGCTG 1133
QY 1140 CACAGTGTCTTGGCGCTGAGGAGCGTGGCGCTCTATGTGGCGCTCATCGCGTGGCGCTG 1199
Db 1134 CACACTTCTTCGGCGCGGAGAGCGTGTCTCTATCATCGCGCTCTGCGCTGGCGCTG 1193
QY 1200 TGCCTGTCTCTGCTGTCTGTCTCATCTCTCGTTTATTTGCCGGAAGAGAGGGGCTG 1259
Db 1194 TGCCTCATCTTGTCTGTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1253
QY 1260 GACTCAGATGTGGTGTGACTGTTCATCTCTACCTCAGGCTTCCAGCCCGTCCAGCATCAAG 1319
Db 1254 GACTCAGAGTGGTGTGACTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1313
QY 1320 CCCAGCAAGCAGACCAACCCCATCTGCTCCACCATCTCCAGCGGAGCCTCAGCACCAACCC 1376
Db 1314 CCCAGCAAGCAGACCAATCCCCCATCTGCTCCACCATCTCCAAACCGGAGCCTCAGCACCA 1373
QY 1377 ACCACCTACCAAGGAGTCTCTCTCCCGGAGGATGGCGCCAGCCAGCCTCAGCACCAACCC 1436
Db 1374 ACCACTTACCAAGGAGTCTCTCTCCCGGAGGATGGCGCCAGCCTCAGCACCAACCCAG 1433
QY 1437 ACCAATGGGCACTGTCTCAGCCCCCTGGTGGCGCGCCACACACTGACACCAAGCTCT 1496
Db 1434 TCTAATGGTCACTGTCTCAGCCCCACTGGGAGTGGCGCCATACGCTGACCAACAGCTCC 1493
QY 1497 CCCACCTCTGAGGCGGAGGAGTGTCTCTCCGCTCTCTCCACCCAGAACTACTTCCGCTCC 1556
Db 1494 CCCACCTCTGAGGCTGAGGAGTGTCTCTCCGCTCTCTCCACCCAGAACTACTTCTCTCT 1553
QY 1557 CTGCCCCGAGGCAACAGCAACATGACCTATGGGAGCTTCACTTCTCTCGGGGCGCGGTG 1616
Db 1554 CTGCCCCGCGGTACCAAGCAACATGGCCCTATGGAGCTTCACTTCTCTCGGGGCGCGGTG 1613
QY 1617 ATGATCTCTAATACAGGTATCAGGCTCTCATCCCCCAGATGCATATCCCCCGAGGGAAG 1676
Db 1614 ATGATCTCTAATACAGGAAATCAGGCTCTCATACCCCCGAGGGAAGTCTCCCCCGAGGGA 1673
QY 1677 ATCTATGAGATCTACCTCAGCTGCACAGCGGAGAGACGTGAGGTGGCCCCCTAGCTGGC 1736
Db 1674 ATCTAAGAGATCTACCTCAGCTGCACAGCGGAGAGACGTGAGGTGGCCCCCTAGCTGGC 1733
QY 1737 TGT CAGACCTGTGTGAGTCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1796
Db 1734 TGT CAGACCTGTGTGAGTCCCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1793
```

1797 CCAGTATCTCTGGCTATGAGACCACTGTGGGAGAGCCAGCCCTGACAGCTGAGCCTGGC 1856  
1794 CCAGTATCTCTGGCTATGAGACCACTGTGGGAGAGCCAGCTGAGCCTGGC 1853  
1857 CTCAAAAAGCAGTCTGTGGAGGAGCTGGGAGCAGGAGTGTCTGCACCTGGCGGAGGAG 1916  
1854 CTCAAAAAGCAGTCTGTGGAGGAGCTGGG---AGGACGTGCTGCACCTGGTGGAGGAG 1910  
1917 GCGCCCTCCACCTCTACTACTGCGAGCTGGAGGCGCAGTGTCTGCTCTTCAACCGAG 1976  
1911 TCGCCTCTCATCTCTACTACTGCGAGCTGGAGGCGGCGCTGTATGTCTTTACCGAG 1970  
1977 CAGTGGGCGCTTTGGCCTGTGGAGAGGCGCTCAGGTGGCTGCGCCGAGCGCTC 2036  
1971 CAGCTAGGCGCTTTGGCCTGTGGAGAGGCGCTCAGGTGGCTGCGCCGAGCGCTC 2030  
2037 AAGCTGCTTCTGTTGGCGCGCTGCTGACCTCCCTCGAGTCAAAATCCGGGTCTAC 2096  
2031 AGGCTCTTCTGTTGGCGCGCTGCTGACCTCCCTCGAGTCAAAATCCGGGTCTAC 2090  
2097 TGCTGCAATGACACCCAGCATCTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2156  
2091 TGCTGCAATGACACCCAGCATCTCAAGAGGTGGTGCAGCTGGAGAGCAGCTGGG 2150  
2157 GGACAGCTGATCAGAGGACCGGTCTGCACTTCAAGGAGTGTACACACCTGGC 2216  
2151 GGACAGCTGATCAGAGGACCGGTCTGCACTTCAAGGAGTGTACACACCTGGC 2210  
2217 CTATCATCATCAGATGTGCGGAGCTCTCTGTGGAGAGTAAAGTCTCTGCTGAGCTACCA 2276  
2211 CTGTCTCATCAGATGTGCGGAGCTCTCTGTGGAGAGTAAAGTCTCTGCTGAGCTACCA 2270  
2277 GAGATCCCTTTTATCACAATCTGGAATGCGAGCGGCTGCTGCACTGCACTTCAAC 2336  
2271 GAGATCCCTTTTATCACAATCTGGAATGCGAGCGGCTGCTGCACTTCAAC 2330  
2337 CTGGAGCGTGTAGCCCGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2396  
2331 CTGGAGCGTGTAGCCCGAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2390  
2397 GAGGCGGAGCGGAGAGCTTCAAGTCACTTCAACATCAACCAAGGAGCAGAGTGTGCT 2456  
2391 GAGGCGGAGCGGAGAGCTTCAAGTCACTTCAACATCACTAAGGAGCAGAGTGTGCT 2450  
2457 GAGCTCTGCTGCTGAGAGTGAAGCGGGGCTCCAGCCCTGTGGGCGCCAGTGCCTTC 2516  
2451 GAAATGCTGCTGCTGAGAGTGAAGCGGGGCTCCAGCCCTGTGGGCGCCAGTGCCTTC 2510  
2517 AAGATCCCTTCTCTATTGGCAGAGATAATTTCCAGCTGGACCCAGCCCTGTAGCGG 2576  
2511 AAGATCCCTTCTCTATTGGCAGAGATAATTTCCAGCTGGACCCAGCCCTGTAGCGG 2570  
2577 GGTGCGGAGTGGCGGACTCTGGCCGAGAACTCCAGCTGGACAGCCATCTCAGCTTCTTT 2636  
2571 GGTGCGGAGTGGCGGACTCTGGCCGAGAACTCCAGCTGGACAGCCATCTCAGCTTCTTT 2630  
2637 GCTTCAAGCCAGCCAGCCAGCAGCAGTCTCTCACTGTGGAGGCGGCGGCTTCCCC 2696  
2631 GCTTCAAGCCAGCCAGCCAGCAGCAGTCTCTCACTGTGGAGGCGGCGGCTTCCCC 2690  
2697 AACGGCAACTCAGCAGCTGGCTGCGAGCAGTGTGCTGCTGCGGAGCAGAGCTGGC 2756  
2691 AACGGCAACTCAGCAGCTGGCTGCGAGCAGTGTGCTGCGGAGCAGAGCTGGC 2750  
2757 CTCTTCAAGTGTGGAGGCTGAGTGTGAGGCGGCGGAGCCGAGCAGCAGTCTT 2814  
2751 CTCTTCAAGTGTGGAGGCTGAGTGTGAGGCGGCGGAGCCGAGCAGCAGTCTT 2810  
2815 CACAGCTTTGGCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2874  
2811 CACAGCTTTGGCAGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2870  
2875 GGGGAGA 2881

Db 2871 CCGGGGA 2877  
RESULT 10  
BC009333  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
REMARK  
COMMENT  
FEATURES  
source  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov  
Series: IRAL Plate: 26 Row: g Column: 22.  
1. .2688  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

BC009333 2688 bp mRNA linear PRI 19-DEC-2003  
Homo sapiens cDNA clone IMAGE:4126760, partial cds.  
BC009333  
BC009333.2 GI:40226527  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2688)  
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.P., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Lohmeier, S., Peters, G.J., Carninci, P., Brance, C., Raha, S.S., Loquellano, N.A., McEwan, P.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, K., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
2 (bases 1 to 2688)  
Strausberg, R.  
Direct Submission  
Submitted (12-JUN-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
NIH-MGC Project URL: http://mgi.nci.nih.gov  
On Dec 19, 2003 this sequence version replaced gi:14424611.  
Contact: MGC help desk  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
Gaithersburg, Maryland;  
Web site: http://www.nisic.nih.gov/  
Contact: nisc.mgc@nih.gov  
Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.D., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lari, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsougeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.





[illegible]





AUTHORS Ackerman,S.L., Kozak,L.P., Przyborski,S.A., Rund,L.A., Boyer,B.B. and Knowles,B.B.  
 TITLE The mouse rostral cerebellar malformation gene encodes an UNC-5-like protein  
 JOURNAL Nature 386 (6627), 838-842 (1997)  
 MEDLINE 97271898  
 PUBMED 9126743  
 REFERENCE 2 (bases 1 to 9299)  
 AUTHORS Ackerman,S.L., Kozak,L.P., Rund,L.A. and Knowles,B.B.  
 TITLE Direct Submission  
 JOURNAL Submitted (25-SEP-1996) The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA  
 FEATURES  
 Location/Qualifiers  
 1..9299  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57B6/SJL"  
 /db\_xref="taxon:10090"  
 /chromosome="3"  
 /map="between Amy1 and Mpmv9"  
 1..9299  
 /gene="rcm"  
 151..2946  
 /gene="rcm"  
 gene  
 CDS  
 /function="transmembrane receptor"  
 /note="similar to C. elegans UNC-5"  
 /codon\_start=1  
 /product="rostral cerebellar malformation protein"  
 /protein\_id="AAB54103.1"  
 /db\_xref="GI:2088527"  
 /translation="MRKGLRATAARCGGLGVLQMLVLPALALLSASGTGSAQODDE  
 FPHLEPFPDPPELPHLEPEAYIVKKNPNVLYKALSPAQIYPKNSWVHQ  
 KDHVDERVDSGLIVREVSIEIRQVEELFGPEYWCQVASSAGTTKSRKAYV  
 RIAYLRKPEQPLEGVELEVLQCRPEGIPVAEVLKNEIDIDPAEDRFYI  
 TIDHNLIIKQRLSDTANYTCAKNIIVAKRKSTATVIVVANGSMSTWDSVNSRC  
 GRGYKRIETCNPLNGAFEGOSVOKIATLCPVDGRWTSWKSMTGCTECLA  
 WRRECTAPAKNGKDCGLVLQSKNDCGLCMQAPDSDVALVGVIVATVCLIA  
 ITVVALFVKNRHFESDIIDSSALNGFPQVNIKAARQDLIAVPDLTAAAMRY  
 GPVALHDVSDKIPTNPSILPLPNLKITVNSGAVTPQDDLAFFSKLSPQMTQS  
 LLEENALINKSLAQTDPSCTAFGTNSLGHLLIPNSGVSLLPAGAI PQGRVYE  
 MYVTVHRKNRPMEDSOTLLTPVSCGPPGALLTRPVILTHHCADSPEDMKIQL  
 KNOVQOGEDVVVVEENFTPCYIQLDAEACHILTENLTVLVQSTTKAAAKRL  
 KLAIFGLCSLSIYSIRVYCLDDTDQALKEVLOROMGOLLEPKALHPFKGLTHN  
 LRLSHDIAHLSWLSKLAKYQEIFYHIWSGQRLNHCFTLERLSINTVELVKLC  
 VRQVEGEQIFQNLCTVSEPTGIDILDPASTITVTGSPAFSLPLPIRKLCSLL  
 DAPQTRGHDRWRLAHKLNLDRLNYEATKSSPTGVILDLWEAQNFDPGNLSMLAAVLE  
 EMGRHETVTVSLAAGQY"

## ORIGIN

Query Match 32.6%; Score 940.2; DB 10; Length 9299;  
 Best Local Similarity 61.9%; Pred. No. 9,9e-146;  
 Matches 1642; Conservative 0; Mismatches 943; Indels 66; Gaps 7;

QY	190	TGCTGGTGCACACCGGACCTGCTCCGCCACTTCCTGGTGGAGCCCGAGAGTGTGTACA	249
Db	308	TTCCATCTGACCACTGAGCCATTGCCACACTTCCTCATTTGAGCCCGAGAGCTTACA	367
QY	250	TCGCTCAAGAACAAAGCCAGTCTGCTGTGTGTGCAAGGCCGTGCCCGCCACAGACTTCT	309
Db	368	TTGTGAAGAACAAAGCCGTGTGAACCTGTATTGTGTAAGCCAGCCCTGCCACCCAGACTTACT	427
QY	310	TCAAGTCAACGGGAGTGGTGGCCGACGAGTGGACACAGCTGATCGAGCGGACGACACG	369
Db	428	TCAAGTCAACGCGAGTGGGTTTCATCAGAAGGACCCACGTTAGTAGAGAGAGTAGATG	487
QY	370	GGAGCAGTGTGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGCAGCAGGTTCGAGA	429
Db	488	AAACCTCTGGTCTTAATTGTGAGAGAAGTGGACATTGAGATTTCAGCCGACGAGTGGAG	547
QY	430	AGGTGTTTGGGCTGGAGGAATATCTGGTCCAGTCCGTTGGCATGGAGTCTCTCGGGACCA	489
Db	548	AACTGTTTGGGCTGAAGATTACTGGTGCCAGTGTGTGGCTGGAGTCTCAGCAGGCACTA	607

QY 490 CCAAGAGTCAGAAGCGCTTACATCCGATAGCCAGATTGCGCAAGAACTTTCAGAGGAGC 549  
 Db 608 CGAAGATCGGAAGGCATACCTGCGCATTTGCTATCTCGGAAGACATTCGAGCAGGAAC 667  
 QY 550 CGCTGGCCAAAGGAGTGTCCCTCGAGCAGGGAATGCTGCTCCCTGCGCTCCACCGGAG 609  
 Db 668 CTTTGGGAAAGGAAGTGTCTTGGAGCAGGAAGTCTTACTCCAGTGTCTGGCCACCTGAAG 727  
 QY 610 GCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGAGACCTGTGTGACCCGCTCCCTGG 669  
 Db 728 GGATCCCAAGTGGCTGAGGTGAATGGCTAAAGAAATGAAGACATAATTTGATCTCTGTGAAG 787  
 QY 670 ACCCCAATGTATACATCACGGGGAGCAGACGCTGTGTGCGACAGGCGCCGCTGTGCTG 729  
 Db 788 ATCGGAACTTTTATATTACTATCGATCACAACTGATCATCAAGCAAGCCGCACTCTCAG 847  
 QY 730 ACACGGCCAACTACACCTGCTGGCCAAAGAACATCTGTGGCAGCTGCGCCGACGGCTCCG 789  
 Db 848 ATACAGCAAAATTATACCTGTGTGGCAAAAATATTGTGCCAAGAGAAAAGCACACAG 907  
 QY 790 CTGCTGTCTCATCTCTACGTGAACGGTGGTGTGCTGACGTGGACCGAGTGGTCTCTCTGCA 849  
 Db 908 CCACCTGTCTCATCTGTATTTAATGTTGCTGGTCCACCTGGACAGAGTGTCTGTGTGA 967  
 QY 850 GCGCCAGCTGTGGCGCGGCTCGCAGAAAACGAGCGGAGCTGCACCAACCCGCGGCGCTC 909  
 Db 968 ACAGCCGCTGTGGCGAGGATATCAGAAACGCAAGAACCTGTGCACCAACCCGAGCCAC 1027  
 QY 910 TCAACGGGGCGCTTCTGTGAGGGGAGAAATGTCCATGACCGCACCGCTCTCTCTCTGC 969  
 Db 1028 TCAATGGTGGGCGCTTCTGTGAGGGGAGAGTGTGCAGAA---AATAGCATGCATACCGT 1084  
 QY 970 TTGCTCTCTGTGACGCGCAGCTCGAGCCCTGGAGCAAGTGGTGGGCTGTGGGCTGCACT 1029  
 Db 1085 TATGTCAGTGTGATGTAGTGCACTTCAATGAGCAAAATGGTCAACTGTGGAGTGAAT 1144  
 QY 1030 GCACCCACTGTGGCGGAGCCCTGTGCTGTGACCCAGCACCCCGCAACGAGAGGGAGAGT 1089  
 Db 1145 GCACCCACTGTGGCGAGGAGGTGTACAGCAGCAGCCGCCCAAGAAACGCGGGTAAAGACT 1204  
 QY 1090 GCACGGGCACTGACTGTGACACCCGCAACTGTACAGTGAACCTCTGTGTACAGTGCCT 1149  
 Db 1205 GTGATGGCTGTCTCTCCAAATCCAAAGAACTGCACTGATGGGTGTGCACTGCAGGCTGCTC 1264  
 QY 1150 CTGGCCCTCAGGACGTCGCTCTATGTGG---CCTCATGCGCGTGGCGCTGTGCTGG 1206  
 Db 1265 CTGACTCAGATGATGTGGCTCTCTACGTGGGATTTGTATCGTGTAAACAGTCTGTCTCG 1324  
 QY 1207 TCCTGCTGTGCTGTCTCATCTCTGTTTATTCGCGAAGAGAGGGGCTGGACTCAG 1266  
 Db 1325 CGATCACTGTGTGGTGGCCCTGTTTGTGTATCGGAAGAACCCCGTGACTTTGAGTCTG 1384  
 QY 1267 ATGTGGTGACTGTCTCATCTTCACTCAGGCTTCAGCCCGTTCAGCATCAAGCCGACGA 1326  
 Db 1385 ACATCATTTGACTCTCTCAGCACTCAATGGGCGCTTTCAGCCTGTGAACATCAAGGCTGCCA 1444  
 QY 1327 AAGCAGACAAACCCCATCTCTCACCATCCAGCGGACCTTCAGCACCACCACTTACC 1386  
 Db 1445 GACAAGATCTCTGGGTGTCCTCCCTGACC-----TCACCTCAGTGCAGCCATGTACA 1498  
 QY 1387 AGGCGAGTCTCTGTCCCGGCGAGGATG-----GGCCGAGCCCAAGTTTCAGCTCACCAC 1441  
 Db 1499 GGGGACCTGTCTATGCTCTGCATGATGTCTCAGACAAAATCCCAATGACCAACTCTCCA 1558  
 QY 1442 TGGGCACTGCTCAGCCCTTGGGTGGCGGCCCAACACACTGCACACAGCTCTCCAC 1501  
 Db 1559 TTTGGAACCACTACCCAACTTGAATCAAAATGTAACAGCTCAGGTGTGTCTACTC 1618  
 QY 1502 CTC-----TGAGGCGGAGGAGTGTGCTCTCCCGCTCTCCACCC----- 1539  
 Db 1619 CTAGGATGACCTTGCCGAGTCTCTCATCCAACTGTCAACCCAGATGACCCAGTCTTTCG 1678  
 QY 1540 -----AGAACTACTTCCGCTCCCTGCCCGGAGGACCAAGCA 1575

Db	1679	TAGAGATGAGCCCTTAACCTGAAGAACACAGAGCCTCGCAAGACAGACTGACCCATCCT	1739
Qy	1576	ACATGACCTATGGGACCTTCACTTCTCGGGGGCCGCTGATGATCCCTTAATACAGGTA	1635
Db	1739	GCAAGCAATTGGTACCTTCAACTCTCTTGGGGGTCACCTCATCATTTCTTAATTCAGGAG	1798
Qy	1636	TAAGCCTCCTCATCCCCCAGATGCATACCCAGGGAAGATCTATGAGATCTPACCTCA	1695
Db	1799	TAAGCTTGCTGATTCGCGCTGGGCCATTCCTCAGGGGAGAGTCTATGAATGTATGTGA	1858
Qy	1696	CGCTGCACAAGCCGGAAGAGTGAAGTTGCCCTAGCTGGGTGTCAAGCCCTGCTGAGTC	1755
Db	1859	CTGTACAAGAAAGAAATATGAGCCCCCATGGAAGACTCTCAGACCTACTTACCC	1918
Qy	1756	CCATCGTTAGCTGTGGACCCCTGCGCTCTGCTCACCGGCCAGTCACTCTGGCTATGG	1815
Db	1919	CTGTGGTGAAGCTGTGGGCTCTCTGGAGCTCTGTGACCCGCCCTGTCTCTCACTCTGC	1978
Qy	1816	ACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCCTCGGCCTCAAAAGACAGTCGTGG	1875
Db	1979	ATCACTGTGCAGAGCCCAAGCACCAGAGACTGGAAGATCCAGCTCAAAACCCAGCAGTC	2038
Qy	1876	AGGCAGCTGGGAGCAGAGATGTGCTGCACCTGGCGGAGGAGGCCCTCCCACTCTACT	1935
Db	2039	AGGACAAATGGAGAGTGTGTGGT--GTTGGGAGGAGAACTTCACAACCCCTGTT	2095
Qy	1936	ACTGCCAGCTGGAGCCAGTGCCTGTACTGCTTTCACCGAGCAGCTGGGCGGTTTGCCC	1995
Db	2096	ACATTCAGCTGGAATGCAGAGCTTTGCCATATCCTCAAGAGACCTCAGTACATATGCC	2155
Qy	1996	TGTTGGGAGGCCCTCAGCGTGCTGCGGCAAGCGCCTCAAGCTGCTTCTGTTGCGC	2055
Db	2156	TGGTTGGCAGTCCACCACCAAGACGCTGCAGCGTCTTAACTGGCCANTTTTGGGC	2215
Qy	2056	CGGTGGCTGCACCTCCCTCGAGTACAACATTCGGGTCTACTGCTGTCATGACACCCAG	2115
Db	2216	CCCTCTGCTGCTCTTCCCTGGAGTACAGCATTAGAGTCTACTGCTGGATGACACAGG	2275
Qy	2116	ATGACTCAAGAGGTGGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGC	2175
Db	2276	ATGCCCTGAAGAGATTCTCAACTGGAGAGGCAATGGGAGCAGCTCCTAGAGAAC	2335
Qy	2176	CACGGGTCTTCGACTCAAGAGCAGTTACACAACCTGGCCTATCGATCCAGATGTGC	2235
Db	2336	CCAGGCTCTCATTTTAAAGGAGCAGTCACAACCTGGCGCTGTATTCATGACATCG	2395
Qy	2236	CCAGCTCCCTGTGGAAGATPAAGCTCCTGTGACTACCAGGAGATCCCTTTATCA	2295
Db	2396	CCATTCCTCTCGAAGAGCAATTTGCTGCTAAGTATCAGGAATTCATTTTACCACA	2455
Qy	2296	TCTGGATGCAAGCAGCGGTACTTGCACTGCACTTCACCTGAGAGGTGTGAGCCCA	2355
Db	2456	TCTGGAGTGGCTCTCAAAAGAAACCTCACTGCACTTCACTCTGGAAGACTCAGCCTAA	2515
Qy	2356	GCCTAGTACCTGGCTGCAAGCTGTGGGTGGCAGGTGGAGGCGCAGCGGSCAGAGCT	2415
Db	2516	ACACAGTGAACCTGGTTTGCAACTCTGTGTGGCGAGGTTGAAGAGAAGGGCAGATCT	2575
Qy	2416	TCAGCATCACTTCAACATCAACAAGGACACAAGGTTTGTCTGAGCTGTGGCTCTGGAGA	2475
Db	2576	TCCAGCTCAACTGTACTGTGTAGAGAGAACCTACTGGCATCGACTTACCTCTCTGGACC	2635
Qy	2476	GTGAAGCGGGGTCCAGACCCCTGGTGGGCCCGAGTGCCTTCAAGATCCCTTCTCATTC	2535
Db	2636	CTGCTAGTACCATCAACACTGTCAACGGACCAAGTGTCTTCAGCATTCCTCTCCCTATCC	2695
Qy	2536	GGCAGAAGATAATTTCCAGCTGGACCCACCTGTAGCGGGGTGCGAGTGGCGGACTC	2595
Db	2696	GGCAGAAGCTATGCAGCAGCCTGGATGCCCCCTCAAAACAGAGGCCATGACTGGAGGATGC	2755
Qy	2596	TGCCCCAGAACTCCACCTGGACAGCCATCTCAGCTTCTTTGGCTTCCAGCCAGCCCA	2655

Db	2756	TGSCCAATAAACTCAACTGCAGGTACTTGAATTACTTTGCCACCAATTCGAGCCAA	2815
QY	2656	CAGCCATGATCCTCAACCTGTGGAGGGCGGCACCTCCCAACGCGAACTCAGCCAGC	2715
Db	2816	CTGGCGTAATCCTGCATCTTTGGAGACACAGACTTCCCAAGATGGAACTCAGCATGC	2875
QY	2716	TGGCTGCAGCAGTGCTGCACTGGGCCAGCAGCGCTGGCGCTCTTACAGTGTGGAGG	2775
Db	2876	TGGCAGCCGCTCTGGAAGAAATGGAGACATGACAGCTGGTGTCTTGGCAGCAGAG	2935
QY	2776	CTGAGTGTCTGA	2786
Db	2936	GACATATTGA	2946

RESULT 14	AY187310	2962 bp	linear	VRT 06-JUN-2003
LOCUS	Gallus gallus UNCS-like protein 3 mRNA, complete cds.			
DEFINITION	AY187310			
ACCESSION	AY187310.1	GI:31442350		
VERSION				
KEYWORDS				
SOURCE	Gallus gallus (chicken)			
ORGANISM	Gallus gallus			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Archosauria; Aves; Neognathae; Galliformes; Phasianidae;			
	Phasianinae; Gallus			
REFERENCE	1 (bases 1 to 2962)			
AUTHORS	Guan, W. and Condict, M.L.			
TITLE	Characterization of Netrin-1, Neogenin and cUNC-5H3 expression during chick dorsal root ganglia development			
JOURNAL	Gene Expr. Patterns 3, 369-373 (2003)			
REFERENCE	2 (bases 1 to 2962)			
AUTHORS	Guan, W. and Condict, M.L.			
TITLE	Direct Submission			
JOURNAL	Submitted (26-NOV-2002) Neurobiology & Anatomy, University of Utah, 20 North, 1900 East, Salt Lake City, UT 84132-3401, USA			
FEATURES	Location/Qualifiers			
source	1. .2962			
	/organism="Gallus gallus"			
	/mol_type="mRNA"			
	/db_xref="taxon:9031"			
CDS	1. .2796			
	/codon_start=1			
	/product="UNCS-like protein 3"			
	/protein_id="AAO67275.1"			
	/db_xref="GI:31442351"			
	/translation="MGKGLGHTAARCGLCGMGYLLHSHVVLPAVALGASRPGSAAQDD FHELETPSPDPEPLPHELLEPEAYIYKKNPNLYCKASPAQIYFKCNSEWHQ KADYDERVDETSGLIVCEVSIIEISHQVEELFGPEDYWCQVCWASSAGTTKSRKAY RIALRKTEPLGKEVSELEQVLLQCPPEGIPIVAEVLWKNKEVIDIPDPRNFI TIDHNI1IKQARLSPTANDTVCANIKIVAKRKTTATVIVVNGSACRNSRRC GRGQFRTCTNPAPLNGCATCEGONVOKIACITLCPVDGKWTSMKSTGCTECTH WRRECTPAPKNGKDCGLVILQSKNCTDGLCMQAAPDSDVALIVGIVIAVICLA ISVVALFVYKXNRHDFESDIIDSNALNGGFPVNIKAARQDLAVPPDLTSAAMAYR GPVALHDVSK1PTMNSIPDLPENLKIKVNTGAVTPQDELSDFSKLSQITVS LUENETVHWNKSLARQSDPTACTFATGNSLGHV1PNSGVSLIIPAGAPQGRVQ MYVTHRKEGMRPPVEDSTLITPVVSCGPALLTRPVLTTHHCAPENMDWQIOL KHQAGQGWEDVWVGEENFTTVCV1QIDPEACHILTETLSTVALVQSITKAAAKRL KLAIFGLSCSLEYIRVYCLDDTQDALKEVLQIRQMGGQLLEPKTLHFPGSTHN LRSLSHDIASHLWKLPAKYQIE1FFYHWSGCGNRLHCTFTLFRFSLNLELVCKL VQVEGEGQ1FQNCVSVEEPTG1TIVGPNFAS1PIPIRQKLCSS1 DAPQTRGHDRWRLAHKLRLYLNIYFATKSSPTGVILDLWEAQFPFDGNSLMAVLE EMGRHETVTVSLAEGNY"			

## ORIGIN

	Query Match	32.6%	Score 938.4	DB 5	Length 2962
	Best Local Similarity	61.3% <td>Pred. No. 2.2e-145</td> <td></td> <td></td>	Pred. No. 2.2e-145		
	Matches 163	Conservative	0	Mismatches 961	Indels 72 Gaps 5
Qy	190	TGCTGGTGCACCGGACCTCTCCCACTTCCTGTGGAGCCGAGATGTGACA	249		
Db	158	TTCTTCTGATCTCCAGAGCAATTCGCCCACTTCTCAATGAACCCGAAGAGCTTACA	217		



```
Db 2363 TCATACCCCTGGAGCTGCTGCTGAAAACCTCTGTGTGCGCAAGTCAAGGAGAGGGCAGA 2422
Qy 2413 GCTTCAGCATCAACTTCAACATCACCAGGACACAAAGTTTGTGAGCTGCTGCTGG 2472
Db 2423 TCTTCAGCTGAACCTGCTAGTATCAGAGGAACCCACTGGCATTTGATATCCCATCATGG 2482
Qy 2473 AGAGTGAAGCGGGGTCACAGCCCTGTGTGGGCCCCAGTGCCTTCAAGATCCCTTCCTCA 2532
Db 2483 ATTACAGCAGCAGCATCACAAAGATAGTTGGGCCCAAGGTTTCAGCATCCCTCCCAA 2542
Qy 2533 TTCCGAGAGAGATAATTTCAGCTGTGAACCCACCTGTAGCGGGGTGCCAGCTGGCGGA 2592
Db 2543 TAAGGAGAGAGCTCTGAGCAGCCTGTGATGCACCCAGACCCCGGGGCCATGACTGGAGGA 2602
Qy 2593 CTCTGGCCCCAGAAACTCCACTGACAGCCATCTCAGCTTCTTTGCTTCAAGCCAGCC 2652
Db 2603 TGTGCCCCACAGCTGAATTGGACAGTAGTACCTAAATTAATTTGTACGAAGTCAGATC 2662
Qy 2653 CCACAGCCATGATCTCAACTGTGGGAGCGGCGACTTCCCAAGCGCAACTCAGCC 2712
Db 2663 CCACTGGGGTGAATCTCTGGATCTCTGGAGAGCCAGAAATTTCCCTGATGGCAACTGAGCA 2722
Qy 2713 AGCTGCTGCAGCAGTGGTGAAGTGGCCAGCAGCAGCGTGGCTCTTTCACAGTGTGG 2772
Db 2723 TGTGCGCAGCAGTGTGGAAGAAATGGAGCAGATGAACCCGTTGTTCTTTGGCAGCAG 2782
Qy 2773 AGCTGAGTGTGAGCGCGGCCAGGCCGA 2802
Db 2783 AAGGAATTAATGATGACAGCTTGGCGTGA 2812
```

## RESULT 15

```
AB118026
LOCUS AB118026 9328 bp mRNA linear ROD 19-DEC-2003
DEFINITION Rattus norvegicus Unc5h3 mRNA, complete cds.
ACCESSION AB118026
VERSION AB118026.1 GI:40217509
KEYWORDS
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

## REFERENCE

1 Kuramoto, T., Kuwamura, M. and Serikawa, T.  
Rat neurological mutations cerebellar vermis defect and hobble are caused by mutations in the neurin-1 receptor gene Unc5h3  
Mol. Brain Res. (2003) In press  
2 (bases 1 to 9328)  
Kuramoto, T. and Serikawa, T.  
Direct Submission  
Submitted (22-AUG-2003) Takashi Kuramoto, Institute of Laboratory Animals, Graduate School of Medicine, Kyoto University;  
Yoshidakone-cho, Sakyo-ku, Kyoto 606-8501, Japan  
(E-mail:tkuramoto@anim.med.kyoto-u.ac.jp,  
URL:www.anim.med.kyoto-u.ac.jp, Tel:81-75-753-4494,  
Fax:81-75-753-4409)

## FEATURES

```
source
1..9328
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/chromosome="2"
1..9328
/gene="Unc5h3"
151..2946
/gene="Unc5h3"
/codon_start=1
/protein_id="BAD05181.1"
/db_xref="GI:40217510"
/translation="MRKGLRATAARCGLGLYLLQMLVLPALALLSASGTGSAQQDDDD
FFHLPETFPDPEPLPHLEPEAYIVKKNPNLYCKASPATQIVFKCNSEWHQ
```

## gene

## CDS

```
3'UTR
2947..9328
EMGRHBTVVSLAAEQY"
/gene="Unc5h3"
8727..8772
/note="simple sequence repeats"
/rpt_unit="gt"
9309..9314
/gene="Unc5h3"
```

## repeat\_region

## polyA\_signal

## ORIGIN

Query Match 32.5%; Score 936.8; DB 10; Length 9328;  
Best Local Similarity 61.8%; Pred. No. 3.6e-145;  
Matches 1631; Conservative 0; Mismatches 937; Indels 72; Gaps 6;

```
Qy 204 CCGGACCTGCTTCCCCACTTCTCTGGTGGAGCCGAGAGTGTGTACATCTCTCAGAACAG 263
Db 322 CTTGAGCCATTACCACTTTCTCATTTGAGCCCGAGAGCTTACATTTGTGAAGAACAG 381
Qy 264 CCAGTGTCTTGTGTGCAAGCCGTCGCCGACGACAGATCTTCTTCAAGTGCACCGG 323
Db 382 CTTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCTACTTCAAGTGCACAGC 441
Qy 324 GAGTGGTGGCGGAGTGGACCACTGTATCGAGCGCAGCAGACGAGGAGAGTGTGTAG 383
Db 442 GAATGGTTTCATCAGAAAGGACCCAGTAGTAGCAGAGAGGTAGATGAACCTCCGGTCTC 501
Qy 384 CCGACCATGGAGGTCGCGATTAAATGCTCAAGCGCAGCAGTGTGTGAGAGTGTTCGGGCTG 443
Db 502 ATTGTGAGAGAGTGTGAGTGTGAGATTTACCCAGCAGGTGGAGAACTGTTCGGGCT 561
Qy 444 GAGGAATCTGTGTGCGAGTGTGCGTGTGCGATGAGACTCTCGGGCACCACCAAGAGTCAAG 503
Db 562 GAAGATTACTGTGCGCAGTGTGTGCGCTGGAGTCTCAGCGGCACTACGAGAGTGTGAG 621
Qy 504 GCCTACATCCGATAGCCAGATTGGCGAAGAACTTCGAGCAGGAGCCGCTGGCCAAAGAG 563
Db 622 GCATACGTGGCATTTGCGTATCTCGGAAGACATTTGAGCAGGAAACCCCTTGGGAAAGAA 681
Qy 564 GTGTCCCTGGAGAGGATCGTGTGCTGCGCTGCCCTCCACCGAGGAGGATCCCTCCAGCC 623
Db 682 GTGTCTTGGAGCAGAGTGTGTACTCAGTCCCGCCACCTGAGGGATCCCAATGGCT 741
Qy 624 GAGTGGAGTGGCTCCGGAACGAGACCTGTGTGACCCGCTCCCTGGAGCCCAATGTATAC 683
Db 742 GAGTGGAGTGGCTAAAGAAATGAAGACATAATTGATCTGTGAGAGATCGGATTTTAT 801
Qy 684 ATCAGCGGAGCAGCAGCTGTGTGCGCAGACAGCCCGCTTGTGTGACAGCGCCAACTAC 743
Db 802 ATCACTATCGATCAACCTGTATCATCAAGCAGGACAGCTCTCAGATACAGCAATAT 861
Qy 744 ACCTGCGTGGCCAAAGAACATCGTGGCAGCTGCCGCGCAGCGCTCCGCTGTGTATCGTC 803
Db 862 ACCTGTGTGCCAAACATTTGTGCCAAGAGAAAGACCAACAGCCACCGTATCGTG 921
Qy 804 TACGTAAACGTTGGTGGTGTGACGTGGACCGAGTGGTCCGCTCTCGAGCCAGCTGTGGG 863
Db 922 TATGTTAATGTTGGTGGTCCACCTGGCAGAAATGGTCTGTGTGTAAACAGCCGCTGTGG 981
Qy 864 CGCGGCTGGCAGAAACGAGCCGAGCTGCACCAACCCGGCGCTCTCTAAACGGGGCGCT 923
```

Db 982 CGAGGATATCAGAAACGACAAACCTGCACCAACCCAGCCCTCAATGGTGGGCC 1041  
QY 924 TTCTGTGAGGGCAGAAATGTCATCAGCCGACCGTCTCTCTCTGTCTCTCTGTGGAC 983  
Db 1042 TTCTGCGAGGGCAGAGGTGACAGAA--AATAGCATCACCAGTATGCCAGTGGAT 1098  
QY 984 GGCAGCTGAGCCCTGTGAGCAGAGTGGTGGCTGTGGCTGGACTGACACCCACTGGGG 1043  
Db 1099 GGCAGCTGAGCTTCTGTGAGCAAAATGGTCAACTGTGGGACTGAGTGACCAACCTGGGG 1158  
QY 1044 AGCCGTGAGTCTCTGACCCAGCACCAGCGAGGGGAGGAGTGCAGGCGCACTGAC 1103  
Db 1159 AGGAGGGAGTGTACAGCAGCAGCCCAAGAGACGGGGTAAGAGCTGTGATGGCTGGTC 1218  
QY 1104 CTGGACACCCGCAACTGTACACAGTGTCTGTGTACACAGTGTCTTCTGGCCCTGAGGAC 1163  
Db 1219 CTCGATCCAGGACTGACCGATGGCTGTGATGACAGGCTGTCTGACTCAGATGAC 1278  
QY 1164 GTGGCCCTCTATGTGGG---CCTCATCGCGGTGGCCGTCTGCTGTCTGTCTGTCTGT 1220  
Db 1279 GTTGTCTCTACGTTGGGATTTGTATCGCTGTAAACCGTTGTCTGGCAATCACTGTGTG 1338  
QY 1221 GTCTCTATCTCTGTTTATTCGCGGAGAGAGGAGGCTGTGACTCAGATGTGGCTGACTCG 1280  
Db 1339 GTGGCCCTGTCTGTGTATCGCAAGAACCCCGTGACTTTTGTGATCTAATCATGTGACTC 1398  
QY 1281 TCATTTCTCACTCTCAGCTTTCAGCCCGTTCAGCATCAAGCCCAAGCAGACAAACCCC 1340  
Db 1399 TCAGCACTCAATGGCGCTTTCAGCTGTGAACTCAAG-----GCTGCCAGACAA 1449  
QY 1341 CATCTGCTCACTATCAGCCGAGCCTCAGCACCAACCACTACAGGGGAGTCTCTGT 1400  
Db 1450 GATCTCTGCGAGTACCCCGGATCTCACCTCAGCTGCGAGCCATGTACAGGGGACCCGTC 1509  
QY 1401 CCCCAGGAGTGGGCCAGCCCAAGTTCAGCTCACCAGTGGGCACTGTCTCAGCCCC 1460  
Db 1510 TATGCTCTCATATGTCTCAGACAAATCCCAATGACCAATCTCAATCTAGACCCA 1569  
QY 1461 CTGGGTGGCGCGCCACACACTGCACACAGCT-----CTCCCACTCTGAG 1508  
Db 1570 CTGCCCAATTTGAATCAAGTGTACACAGCTCAGGTGCTGTCTCACTCTCAGGACGAC 1629  
QY 1509 GCGGAGAGTGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGGAGGC 1568  
Db 1630 CTGTGCTGAGTCTCATCCAAGCTGTACCCCGAGATGACCAATCTTGTGTAGAGATGAG 1689  
QY 1569 ACCAGCAACATGA-----CCTAT 1586  
Db 1690 GCCTGAACTGAAACAGAGTCTCGAAGACAGACTGACCCATCCTGCAAGCGTTT 1749  
QY 1587 GGGACCTTCAACTTCTCGGGGCGGCTGATATCCCTAATACAGGTATCAGCTTCTC 1646  
Db 1750 GGCACCTTCACTCCCTGGCGGTCACTCATCATCCCTAATCAGGAGTAAGCTTCTG 1809  
QY 1647 ATCCCCCAGATGCCATACCCGAGGAGAGATCTATGAGTCTACCTCAGCTGCACAG 1706  
Db 1810 ATTCCCGCTGGGGCCATTCTCTCAGGGAGAGTCTATGAATGTATGTGCTGTACACAG 1869  
QY 1707 CCGGAGAGCTGAGGTTGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCTGTTAGC 1766  
Db 1870 AAGAAATATGAGCCCTCCATGGAAGACTCTCAGACCTGTGACCCCTGTGTGAGC 1929  
QY 1767 TGTGGACCCCTGGCGTCTGCTCACCAGGCGAGTCACTCTGGTATGGACCACTGTGGG 1826  
Db 1930 TGTGGGCTCTCTGGAGCTCTGTTGACCGCTCTCTCATCTCTCTCACTCTGCTACTGTGCA 1989  
QY 1827 GAGCCAGCCCTGACAGTGTGAGCTGTGCGCTCAAAAGAGAGTGTGCGAGGGGAGCTGG 1886  
Db 1990 GACCTTAACACCGAGGACTGGAAGATCCAGCTCAAAACACGAGGAGTGTGCGGAGCAGTGG 2049  
QY 1887 GAGCAGGATGTGTGCTCCTCTGGGGAGGAGGCGCCCTCCACCTCTACTCTGCGAGCTG 1946

2050 GAGGATGTTGGTTG---TTGGAGAGGAAACATTCAAAACCCCTGTTTACATTGAGCTG 2106  
1947 GAGGCCAGTGCCTGCTAGTCTTACCCAGAGAGCTGGCGCTTTCCTGCTGTGGAGAG 2006  
2107 GATGAGAGGCTTGCATATTTCTCAGGAGAACCTCAGTACATATGCTTGGTGGGCG 2166  
2007 GGCCTCAGCTGGCTGCGCCCAAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCCCTG 2066  
2167 TCACCAACCAAGAGAGCTGCCAAGGCTGTGAGCTGGCCATCTTTGGGCCCCCTCTGCTG 2226  
2067 ACCTCCCTCGAGTACAAACATCCGGGTCTACTGCTGCTGATGACACCAAGATGACTCAG 2126  
2227 TCTTCCCTGGAGTATAGCATCAGAGTCTACTGCTGCTGATGACACAGAGTACCTTGAAG 2286  
2127 GAGTGTGTGACGCTGGAGNAGAGCTGGGGGACAGCTGATCCAGAGGACACGGGTCTG 2186  
2287 GAAGTTCTACAACTGGAAGGCAATGGGAGGACAGCTCTTAAAGAACCCAGGCTCTT 2346  
2187 CACTTCAAGGACAGTTTACCAACCTGGCCCTATCCATCCAGATGTGCCAGCTCCCTG 2246  
2347 CATTTTAAAGGAGCATCCAAACCTGGCCCTCTCAATTCATGACATCACCCATTCCTC 2406  
2247 TGGAGAGTAAGCTCTGTGCTAGCTACAGAGAGTACCCCTTTTATCAGATCTGGAATGCG 2306  
2407 TGGAGAGCAAAATGCTGCTAAGTATCAGGAAATTCATTTTACCACATCTGGAGTGGC 2466  
2307 AGCAGCGTACTGTGACCTGACCTTACCCTGGAGCGTGTGAGCCCGACGACTAGTGAC 2366  
2467 TCTCAAAGAAACCTCCTCCTGTACCTTCTCTCTGGAAGAGTCTGAGCTTAAACAGTAGAG 2526  
2367 CTGGCTTCAAGCTGTGGGTGTGGCAGTGTGGAGGCGACGGGAGAGTCTGAGCATCAAC 2426  
2527 CTGTTTTCAGACTCTGTGCGGAGGTTGAAGGAGAGGCGAGATCTTCCAGCTCAAC 2586  
2427 TTCAACATCACAAGGACACAGAGTTTGTGCTGAGTGTCTGCTGCTGAGAGTGTGAGCGGG 2486  
2587 TGTACTGTGTGAGGAGAACTACTGTGTATCGATTTTGGCTCTCTGAGACCTGTAGCACC 2646  
2487 GTCCAGCCCTGTGGGGCCCGAGTGTCTCAAGATCCCTTCTCTCATTCGGCAGAGAGATA 2546  
2647 ATTACCACTGTCTGAGCAAGTGTCTTTCAGCATCCCTTCTCTATTCGGGAGAGCTA 2706  
2547 ATTTCAGCTGTGACCCACCTCTGTAGCGGGGTGCGGAGTGTGGCGGACTGTGGCCCGAGAAA 2606  
2707 TGTAGCAGCTGGATGCCCCCAACAGAGGCGCATGACTGGAGGATGTCTGGGCCACAAA 2766  
2607 CTCACCTGGAAGCCATCTCAGCTTCTTGTCTTCAAGCCCGAGCCCAAGCCATGATC 2666  
2767 CTTAACCTGAGCAGTACTTGAATTTTGGCCCAAAATCGAGCCCGACTGGCGTAAATC 2826  
2667 CTCACCTGTGGAGGCGGCGACTTCCCAACCGCAACCTCAGCCAGCTGGGTGACGCA 2726  
2827 CTGGATCTCTGGGAGGACAGAACTTCCCGATGGAACCTGAGCATGCTGGCAGCGCTC 2886  
2727 GTGGCTGAGTGTGGCGGAGCAGAGCTGGCTCTTTCAGTGTGCGAGGCTGAGTGTGTA 2786  
2887 CTGGAAGAAATGGGAAGACAGAAACCGTAGTGTCTTCTTGGCAGCAGAGGCTGAGTATGA 2946

Search completed: October 5, 2004, 18:44:04  
Job time : 11157 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 13:21:05 ; Search time 1076 Seconds

(without alignments)  
11374.590 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccgggtctgag.....ccttccccacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq\_29Jan04:\*

- 1: geneseqn1980s:\*
- 2: geneseqn1990s:\*
- 3: geneseqn2000s:\*
- 4: geneseqn2001as:\*
- 5: geneseqn2001bs:\*
- 6: geneseqn2002s:\*
- 7: geneseqn2003as:\*
- 8: geneseqn2003bs:\*
- 9: geneseqn2003cs:\*
- 10: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2881	100.0	2881	6	ABK49422
2	2676.4	92.9	2752	6	ABK37922
3	2615	90.8	2697	6	ABK52891
4	2402.6	83.4	3580	6	ABK15169
5	2296.2	79.7	2907	4	AAK52261
6	2235.8	77.6	3014	2	AAV52940
7	2189.4	76.0	2697	6	AAK16843
8	1570.4	54.5	1787	2	AAV52941
9	1291.2	44.8	9700	4	AAK90958
10	1193.6	41.4	1321	4	AAH99530
11	916	31.8	3646	5	AAH57538
12	889	30.9	2860	6	ABT06279
13	887.4	30.8	2860	6	ABT06280
14	875.8	30.4	2995	6	ABK92062
15	866.4	30.1	2895	6	AAQ93898
16	853.6	29.6	3884	4	AAQ93898
17	853.6	29.6	3884	7	AA5211316
18	853.6	29.6	3884	7	ACD23925
19	853.6	29.6	3884	7	ACA67066
20	853.6	29.6	3884	7	ACA03675
21	853.6	29.6	3884	7	ABX89213
22	853.6	29.6	3884	7	ACD41867
23	853.6	29.6	3884	8	ACA04096
					ADA45664

24	853.6	29.6	3884	8	ADA76095	Ada76095 Human PRO
25	853.6	29.6	3884	8	ADA18745	Ada18745 Human PRO
26	853.6	29.6	3884	8	ADA61368	Ada61368 Homo sapi
27	853.6	29.6	3884	8	ADB19153	ADB19153 Novel hum
28	853.6	29.6	3884	8	ADB27694	ADB27694 cDNA enco
29	853.6	29.6	3884	8	ADA86173	ADA86173 Novel hum
30	853.6	29.6	3884	8	ADB15737	ADB15737 Human PRO
31	853.6	29.6	3884	8	ADA47523	Ada47523 Human PRO
32	853.6	29.6	3884	8	ADA67318	Ada67318 Human PRO
33	853.6	29.6	3884	8	ADB30325	ADB30325 cDNA enco
34	853.6	29.6	3884	8	ADA85621	Ada85621 Novel hum
35	853.6	29.6	3884	8	ADA96833	Ada96833 Human PRO
36	853.6	29.6	3884	8	ADA79137	Ada79137 Human PRO
37	853.6	29.6	3884	8	ADA87276	Ada87276 Novel hum
38	853.6	29.6	3884	8	ADA91570	Ada91570 Novel hum
39	853.6	29.6	3884	8	ADB16478	ADB16478 Human PRO
40	853.6	29.6	3884	8	ADB14633	ADB14633 Human PRO
41	853.6	29.6	3884	8	ADB18594	ADB18594 Novel hum
42	853.6	29.6	3884	8	ADA93809	Ada93809 Human PRO
43	853.6	29.6	3884	8	ADB19705	ADB19705 Novel hum
44	853.6	29.6	3884	8	ADB13017	ADB13017 Human PRO
45	853.6	29.6	3884	8	ACD98496	ACD98496 Novel hum

## ALIGNMENTS

RESULT 1  
ID ABK49422 standard; DNA; 2881 BP.  
XX  
AC ABK49422;  
XX  
DT 15-JUL-2002 (first entry)  
XX  
DE DNA encoding human UNC5-like protein NOV1.

KW Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;  
KW cell signal processing; metabolic pathway modulation; cancerous tissue;  
KW antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;  
KW chromosome 13; gene; ds.  
XX  
OS Homo sapiens.

FH Key Location/Qualifiers  
FT CDS 87..2786  
FT /tag= a  
FT /product= "Human UNC5-like protein NOV1"

XX WO200229038-A2.

PD 11-APR-2002.

PF 04-OCT-2001; 2001WO-US031377.

PR 04-OCT-2000; 2000US-0237862P.

PA (CURA-) CURAGEN CORP.

PI Herrmann JL, Rastelli L, Shimkets RA;

DR WPI; 2002-340104/37.

DR P-ESDB; AAU79939.

XX Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for treating cardiomyopathy, atherosclerosis, and cancer.

PS Claim 8; Page 7-8; 180pp; English.

XX The present invention relates to a new NOVX polypeptide having a 900  
CC (NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 331 (NOV6)  
CC residue amino acid sequence, as given in the specification. The novel  
CC polypeptide, and its encoding polynucleotide, are used to treat



CC cardiomyopathy, atherosclerosis, cancer or a disease related to cell  
CC signal processing and metabolic pathway modulation, in a human. Detecting  
CC the polypeptide or polynucleotide is useful for identifying cancerous  
CC tissue. The antibody can be used to treat diabetes or cancer. The host  
CC cells can be used to produce non-human transgenic animals useful in drug  
CC screening. The present nucleic acid sequence is that of the human UNC5-  
CC like NOV1 gene located on chromosome 13. This sequence encodes the human  
XX UNC5-like protein NOV1 of the invention

SQ Sequence 2881 BP; 526 A; 985 C; 868 G; 502 T; 0 U; 0 Other;

Query Match 100.0%; Score 2881; DB 6; Length 2881;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2881; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGGGGGCCCCGGCCCG 60  
DB 1 AGCTGGGGCTCCGGGCTGAGGCGCTAAAGCCGCCCTCCCGCCCGGGGGCCCCGGCCCG 60

QY 61 GCCCGCCCGCTCCGGGCGCCCGGCGCATGGCCGTCGGGCCCGGCTGTGGCCAGCGCTCC 120  
DB 61 GCCCGCCCGCTCCGGGCGCCCGGCGCATGGCCGTCGGGCCCGGCTGTGGCCAGCGCTCC 120

QY 121 TGGGCATAGTCTCGCGCTTGCTCCGCGCTCGGGTCCCGAGAGTGCACCGTGG 180  
DB 121 TGGGCATAGTCTCGCGCTTGCTCCGCGCTCGGGTCCCGAGAGTGCACCGTGG 180

QY 181 CCAACCCAGTGTGGTGCACACCGGACCTGCTCCCACTTCTGGTGGAGCCCGAGG 240  
DB 181 CCAACCCAGTGTGGTGCACACCGGACCTGCTCCCACTTCTGGTGGAGCCCGAGG 240

QY 241 ATGTGTACATGTCAGAACCAAGCCAGTGTGTTGTGTGCAAGGCCGTGCCCGCACGC 300  
DB 241 ATGTGTACATGTCAGAACCAAGCCAGTGTGTTGTGTGCAAGGCCGTGCCCGCACGC 300

QY 301 AGATCTTCTCAAGTCAACGGGAGTGGTGGCGCCAGGTGGACACAGTATCGAGCGCA 360  
DB 301 AGATCTTCTCAAGTCAACGGGAGTGGTGGCGCCAGGTGGACACAGTATCGAGCGCA 360

QY 361 GCACAGCGGGAGCAGTGTGAGCCGACCATGAGGTCGCAATTAATGTCTCAAGCAGC 420  
DB 361 GCACAGCGGGAGCAGTGTGAGCCGACCATGAGGTCGCAATTAATGTCTCAAGCAGC 420

QY 421 AGTTCAGAGAGTGTTCGGGCTGGAGGAATACGTGTGCCAGTGGCATGAGTCCT 480  
DB 421 AGTTCAGAGAGTGTTCGGGCTGGAGGAATACGTGTGCCAGTGGCATGAGTCCT 480

QY 481 CGGGCACCAACAGAGTCAGAGGCTTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540  
DB 481 CGGGCACCAACAGAGTCAGAGGCTTACATCCGCATAGCCAGATTGCGCAAGAACTTCG 540

QY 541 AGCAGGAGCGCTGGCCAAAGAGGTGTCCTGGAGCAGGCGCATCGTGTGCCCTGCGGTC 600  
DB 541 AGCAGGAGCGCTGGCCAAAGAGGTGTCCTGGAGCAGGCGCATCGTGTGCCCTGCGGTC 600

QY 601 CACCGAGGCGATCCCTCCAGCCGAGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660  
DB 601 CACCGAGGCGATCCCTCCAGCCGAGTGGAGTGGCTCCGGAACGAGGACCTGGTGGACC 660

QY 661 CGTCCCTGGACCCCAATGTATACATACGCGGGAGCACAGCCTGGTGGTCGACAGGCC 720  
DB 661 CGTCCCTGGACCCCAATGTATACATACGCGGGAGCACAGCCTGGTGGTCGACAGGCC 720

QY 721 GCCTTGTGACACGGGCCAACTACACTGTGCGCAAGAAATCATGTGGGACGTGCGCGCA 780  
DB 721 GCCTTGTGACACGGGCCAACTACACTGTGCGCAAGAAATCATGTGGGACGTGCGCGCA 780

QY 781 GCGCCTCCGCTGTGTCTCATGTGAAACGTTGGTGTGACCTGACCGAGTGGT 840  
DB 781 GCGCCTCCGCTGTGTCTCATGTGAAACGTTGGTGTGACCTGACCGAGTGGT 840

QY 841 CCGTCTCAGCGCCAGCTGTGGGCGCGCTGGCAGAACGAGCCGAGGCTGCACCAACC 900  
DB 841 CCGTCTCAGCGCCAGCTGTGGGCGCGCTGGCAGAACGAGCCGAGGCTGCACCAACC 900

DB 841 CCGTCTCAGCGCCAGCTGTGGGCGCGCTGGCAGAACGAGCCGAGCTGCACCAACC 900

QY 901 CGGCGCCTCTCAACCGGGGGCGCTTCTGTGAGGGGCGAGAATCTCCATGACCGACCGTCT 960  
DB 901 CGGCGCCTCTCTCAACCGGGGGCGCTTCTGTGAGGGGCGAGAATCTCCATGACCGACCGTCT 960

QY 961 CCTCTCTCTGTCTCTGTGACGCGCAGCTGGAGCCCGTGGAGCAAGTGTGGCGCTGTS 1020  
DB 961 CCTCTCTCTGTCTCTGTGACGCGCAGCTGGAGCCCGTGGAGCAAGTGTGGCGCTGTS 1020

QY 1021 GSGTGGACTGCACCCACCTGGGAGGACCGTGTGAGTGTCTGACCCAGGACCCCGCAACGAG 1080  
DB 1021 GSGTGGACTGCACCCACCTGGGAGGACCGTGTGAGTGTCTGACCCAGGACCCCGCAACGAG 1080

QY 1081 GGGAGAGTGCACGGGCACTGACCTGGACACCCGCAACTGTACCACTGACCTGTGTATAC 1140  
DB 1081 GGGAGAGTGCACGGGCACTGACCTGGACACCCGCAACTGTACCACTGACCTGTGTATAC 1140

QY 1141 ACAGTGTCTTCTGGCCCTGAGGACGCTGGGCCCTCTATGTGGGCCCTCATGCGCGTGGCGTCT 1200  
DB 1141 ACAGTGTCTTCTGGCCCTGAGGACGCTGGGCCCTCTATGTGGGCCCTCATGCGCGTGGCGTCT 1200

QY 1201 GCGTGTCTGT 1260  
DB 1201 GCGTGTCTGT 1260

QY 1261 ACTCAGATGTGGCTGACCTGCTCCATCTCAGCTCAGGCTTCAGGCCCGTCAGCATCAAGC 1320  
DB 1261 ACTCAGATGTGGCTGACCTGCTCCATCTCAGCTCAGGCTTCAGGCCCGTCAGCATCAAGC 1320

QY 1321 CCAGCAAAAGCAGACAAACCCCATCTGCTCACCATCCAGCCGAGACCTCAGCACCAACCA 1380  
DB 1321 CCAGCAAAAGCAGACAAACCCCATCTGCTCACCATCCAGCCGAGACCTCAGCACCAACCA 1380

QY 1381 CCTACAGGCGAGTCTCTGTCCCGGAGGATGGGCCCAAGCCCAAGTTCAGCTCAGCA 1440  
DB 1381 CCTACAGGCGAGTCTCTGTCCCGGAGGATGGGCCCAAGCCCAAGTTCAGCTCAGCA 1440

QY 1441 ATGGGCACTGTCTCAGCCCGCTGGGTGGCGGCCACACACTGTGCACACAGCTCTCCCA 1500  
DB 1441 ATGGGCACTGTCTCAGCCCGCTGGGTGGCGGCCACACACTGTGCACACAGCTCTCCCA 1500

QY 1501 CCTCTCAGGCGGAGGAGTTCGTCTCCCGCTCTCTCCACCAGAACTACTTCCGCTCCCTGC 1560  
DB 1501 CCTCTCAGGCGGAGGAGTTCGTCTCCCGCTCTCTCCACCAGAACTACTTCCGCTCCCTGC 1560

QY 1561 CCGAGGCGACAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGA 1620  
DB 1561 CCGAGGCGACAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGCGCTGATGA 1620

QY 1621 TCCCTAATACAGGTATCAGCTCTCTATCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680  
DB 1621 TCCCTAATACAGGTATCAGCTCTCTATCCCCCAGATGCCATACCCCGAGGGAAGATCT 1680

QY 1681 ATGAGATCTACCTCAGCTGCACAAAGCCGGAAGAGTGTGGCTGTGCTGTGCTGTGCT 1740  
DB 1681 ATGAGATCTACCTCAGCTGCACAAAGCCGGAAGAGTGTGGCTGTGCTGTGCTGTGCT 1740

QY 1741 AGACCTGTGTAGTCCCATCTGTGTGGAACCCCTGGCGCTCTCAGCCGCGCAG 1800  
DB 1741 AGACCTGTGTAGTCCCATCTGTGTGGAACCCCTGGCGCTCTCAGCCGCGCAG 1800

QY 1801 TCATCTCTGGTATGAGCACTGTGGGAGGCCAGCCCTGACAGCTGGAGGCTGCGCTCA 1860  
DB 1801 TCATCTCTGGTATGAGCACTGTGGGAGGCCAGCCCTGACAGCTGGAGGCTGCGCTCA 1860

QY 1861 AAAAGCAGTCTGCGAGGCGAGTGGAGCAGGATGTGTGTGACCTGGGCGAGGAGCGC 1920  
DB 1861 AAAAGCAGTCTGCGAGGCGAGTGGAGCAGGATGTGTGTGACCTGGGCGAGGAGCGC 1920

QY 1921 CCTTCCACCTCTACTACTGCGAGCTGGAGGCCAGTGTGCTGTCTTCACTCCAGCAGC 1980  
DB 1921 CCTTCCACCTCTACTACTGCGAGCTGGAGGCCAGTGTGCTGTCTTCACTCCAGCAGC 1980

Qy	1981	TGGCCGCTTTTGCCCTGTGGAGAGGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGC	2040
Db	1981	TGGCCGCTTTTGCCCTGTGGAGAGGCCCTCAGCGTGGCTGCCGCAAGCGCTCAAGC	2040
Qy	2041	TGCTTCGTGTTTGGCCGCTGCACCTCCCTCGAGTAAACAATCCGGGTCTACTGCC	2100
Db	2041	TGCTTCGTGTTTGGCCGCTGCACCTCCCTCGAGTAAACAATCCGGGTCTACTGCC	2100
Qy	2101	TGCATGACACCACGATGSCACTCAAGGAGTGGTGCAGCTGGAGAAGAGCTGGGGGAG	2160
Db	2101	TGCATGACACCACGATGSCACTCAAGGAGTGGTGCAGCTGGAGAAGAGCTGGGGGAG	2160
Qy	2161	AGCTGATCCAGGAGCACGGGTCTCGACTTCAAGGACAGTTACCAAACCTCGCCCTAT	2220
Db	2161	AGCTGATCCAGGAGCACGGGTCTCGACTTCAAGGACAGTTACCAAACCTCGCCCTAT	2220
Qy	2221	CCATCCAGGATGTCGACAGCTCCCTGTGGGAAGATAAGTCTCTTGTGAGTACACGAGA	2280
Db	2221	CCATCCAGGATGTCGACAGCTCCCTGTGGGAAGATAAGTCTCTTGTGAGTACACGAGA	2280
Qy	2281	TCCCCTTTTATCATCTGGAATCGCACGACGAGTACTTGCACTGCACCTTCAACCTGG	2340
Db	2281	TCCCCTTTTATCATCTGGAATCGCACGACGAGTACTTGCACTGCACCTTCAACCTGG	2340
Qy	2341	AGCGTGTCAAGCCCCAGCACTAGTAGCTGAGCTGAGCTGAGCTGGGTGCGAGTGGAGG	2400
Db	2341	AGCGTGTCAAGCCCCAGCACTAGTAGCTGAGCTGAGCTGAGCTGGGTGCGAGTGGAGG	2400
Qy	2401	GCGACGGCGAGAGCTTCAGCATCAACTTCAACATCACCAGGACACAAGTTTGTCTGAGC	2460
Db	2401	GCGACGGCGAGAGCTTCAGCATCAACTTCAACATCACCAGGACACAAGTTTGTCTGAGC	2460
Qy	2461	TGCTGGCTCTGGAGAGTGAACGGGGGTCCAGCCCTGTGGGCCCGCAGTGCCTTCAAGA	2520
Db	2461	TGCTGGCTCTGGAGAGTGAACGGGGGTCCAGCCCTGTGGGCCCGCAGTGCCTTCAAGA	2520
Qy	2521	TCCCTTTCCTCATTTGGCGAGAAGATAATTCCAGCCTGGACCCACCTGTAGCGGGGTG	2580
Db	2521	TCCCTTTCCTCATTTGGCGAGAAGATAATTCCAGCCTGGACCCACCTGTAGCGGGGTG	2580
Qy	2581	CCGACTGGCGGACTCTGGGCCAGAACTCCACTGGACAGCCATCTCAGCTTCTTTTGCT	2640
Db	2581	CCGACTGGCGGACTCTGGGCCAGAACTCCACTGGACAGCCATCTCAGCTTCTTTTGCT	2640
Qy	2641	CCAAGCCAGCCCCACAGCCATCATCTCAACTGTGGAGGGCGGCACTTCCCAACG	2700
Db	2641	CCAAGCCAGCCCCACAGCCATCATCTCAACTGTGGAGGGCGGCACTTCCCAACG	2700
Qy	2701	GCAACTTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCGAGCCAGACGCTGGCTCT	2760
Db	2701	GCAACTTCAGCCAGCTGGCTGCAGCAGTGGCTGGACTGGGCGAGCCAGACGCTGGCTCT	2760
Qy	2761	TCAAGTGTGCGAGGCTGAGTGTGAGGCGGGCCAGGCCCGAAGTAAGTCTCAACAG	2820
Db	2761	TCAAGTGTGCGAGGCTGAGTGTGAGGCGGGCCAGGCCCGAAGTAAGTCTCAACAG	2820
Qy	2821	CTTTGGCACCCACCAAGGACAGGACAGCGGAGAGGGGCGCTTCCCCACACCGGGAG	2880
Db	2821	CTTTGGCACCCACCAAGGACAGGACAGCGGAGAGGGGCGCTTCCCCACACCGGGAG	2880
Qy	2881	A 2881	
Db	2881	A 2881	

## RESULT, T 2

RESOLU 2  
ABK37922

ID ABK37922 standard; cDNA; 2752 BP.

XX

AC  
ABK

cdDNA encoding Human protein NOV1.

Human; NOVX; ss; gene; cardiomyopathy; atherosclerosis; diabetes; cell signal processing disorder; metabolic disorder; obesity; infection; anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder; Alzheimer's disease; Parkinson's disease; immune disorder; haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension; osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris; myocardial infarction; ulcer; allergy; benign prostatic hypertrophy; psychosis; neurological disorder; anxiety; schizophrenia; manic depression; dementia; dyskinesia; Huntington's disease; Gilles de la Tourette's syndrome; gene therapy.

**Homo sapiens.**

WO200210216-A2.

07-FEB-2002

30-JUL-2001: 2001WO-US024225.

28-JUL-2000: 2000US-0221409P

20 OCT 2000; 200003-02Z1409F;  
04-AUG-2000; 2000US-0222840P;

04-AUG-2000; 2000US-0223752P.

04-AUG-2000; 2000US-0223762P.

04-AUG-2000; 2000US-0223769P.

04-AUG-2000; 2000US-0223770P.

14-AUG-2000; 2000US-0225146P.  
15-AUG-2000: 2000US-0235393P

13-AUG-2000; 2000US-0225392P.  
15-AUG-2000: 2000US-0225470P

IS-HQ-2000; 2000US-0223470P;  
16-AUG-2000; 2000US-0225697P;

01-FEB-2001; 2001US-0263662P;

05-APR-2001; 2001US-0281645P.

1000

Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;

[illegible]

P-PSDB; AAU85403.

Now, i do not

polypeptide, useful for treating cardiomyopathy, atherosclerosis, infections, cancer, neurodegenerative, metabolic, hematopoietic and immune disorders.

Claim 9; Page 9-10; 213pp; English.

The invention relates to an isolated cytoplasmic, nuclear, membrane bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature form. Also included are the nucleic acids encoding the NOVX proteins, a vector comprising the nucleic acid, a cell comprising the vector, an anti-NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the antibody are useful for treating or preventing a NOVX-associated disorder, where the disorder is selected from cardiomyopathy, atherosclerosis, diabetes, a disorder related to cell signal processing and metabolic pathway modulation, metabolic disorders, obesity, infectious disease, anorexia, cancer-associated cachexia, cancer, neurodegenerative disorders, Alzheimer's disease, Parkinson's disease, immune disorders, haematopoietic disorders, and the various dyslipidaemias, metabolic disturbances associated with obesity, the metabolic syndrome X and wasting disorders associated with chronic diseases, bacterial, fungal, protozoal and viral infections, pain, bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's disease, multiple sclerosis, Albright Hereditary Osteodystrophy, angina pectoris, myocardial infarction, ulcer, allergy, benign prostatic hypertrophy, and psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, and dyskinesias, such as Huntington's disease and Gilles de la Tourette's syndrome. The nucleic acid is useful in gene therapy. The present sequence encodes a NOVX protein

[illegible]













Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorder; arthritis; inflammation. ss.

Homo sapiens

WO200157190-A2.

09-AUG-2001.

05-FEB-2001: 2001WO-US004098

03-FEB-2000: 2000TS-00496914

27-APR-2000; 2000US-00560875.  
20-JUN-2000; 2000US-00560875

19-JUL-2000; 2000US-00620325.

15-SEP-2000; 2000US-00663561.

30-NOV-2000; 2000US-00728422.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT

Ma Y, Zhao QA, Wang D, Wang  
Xue AJ, Yang Y Weibymen T

WPT: 2001 470003/P

P-PSDB; AAM79128.

Nucleic acids encoding

### III DIAGNOSIS AND TREATMENT

CLAIM 1; Page 2691-27

The invention relates to encoded polymers

cytokine, cell proliferation or cell differentiation or which may induce the production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polynucleotides have various structures like nucleic acids.

e.g. stem cell growth factor activity, ha-

activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666 (AAM00020) are omitted as the relevant pages from the sequence listing were missing at the time of publication

Sequence 2907 BP; 552 A; 966 C; 881 G; 508 T; 0 U; 0 Other;

Query Match	79.7%;	Score 2236.2;	DB 4;	Length 2907;
Local Similarity	92.0%;	Pred. No. 0;		
Indels 2565;	Conservative	0;	Mismatches 38;	Indels 186;
Gaps 6				

133 TCGCCGCTTGGCTCCGGGCTCGGGTGCCAGAGTGCCACCGTGGCCACCCAGTGC 192

170 TGGCGGCAGACAGCACGGATGGAGCGCCAGCAGAGTGCCACCGTGGCCACCCACCTCC 338

193 CTGGTGCCAAACCGGACCCTGCTTCCCCCACTTCCTCCTGCAGCCATCATGCCAGTC

[illegible][illegible]

312

TCTTGGGAGGCGGACGAGATCTTCITCA 349

513 AGTCCATCGGGAGTGGGTCGCGCAGGTGGACCACTGATCGAGCGCAGCACAGACGGGA 372

350 AGTGCACCGGGAGTGGGTGCGCCAGGTGGACCACTGATCGAGCGCAGCACAGACGGA 409

Db	410	GCAGTGGGCTGCCCAACATGGAGGTCGGCAATTAATGTCTCAAGGCACGAGGTCGAGAGG	469
QY	433	TGTTTCGGGCTGGAGGAATACTTGTTGGCCAGTGGCGTGGCATGGAGCTCTCTCGGSCACCCACA	492
Db	470	TGTTTCGGGCTGGAGGAATACTTGGTCCCATGTCGTGGCATGGAGCTCTCTCGGSCACCCACA	529
QY	493	AGAGTCAGAAAGGCCTACATCCGATAGCCAGATTGGCAAGAACTTCGAGCAGGAGCCGC	552
Db	530	AGAGTCAGAAAGGCCTACATCCGATAGCCATATTTGGCAAGAACTTCGAGCAGGAGCCGC	589
QY	553	TGSCCAAGAGGTGTCCTCGAGCAGGGCATCGTCTGCTGCCCTGCTCCGCTCAACCGGAGGCA	612
Db	590	TGSCCAAGAGGTGTCCTCGAGCAGGGCATCGTCTGCTGCCCTGCTCCGCTCAACCGGAGGCA	649
QY	613	TCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGCTGGACCCGCTCCCTGGACC	672
Db	650	TCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGCTGGACCCGCTCCCTGGACC	709
QY	673	CCAAATGATATACACGCGGGAGACAGCCTGGTGGTCCGACAGGCCCGCCCTTGCTGACA	732
Db	710	CCAAATGATATACACGCGGGAGACAGCCTGGTGGTCCGACAGGCCCGCCCTTGCTGACA	769
QY	733	CGSCCAACTACACCTCGGTGGCCAGAACATCGTGGCAGCTGCGCGCAGCGCCTCCGCTG	792
Db	770	CGSCCAACTACACCTCGGTGGCCAGAACATCGTGGCAGCTGCGCGCAGCGCCTCCGCTG	829
QY	793	CTGTCAATCGTCTAC-----	806
Db	830	CTGTCAATCGTCTACCGTGGGCCCCGGGACTCCCTGGTCAAGGAGAGGCACTCGCGTGC	889
QY	807	-----GTCAAAGGTGGGTGGTTCGA	825
Db	890	CCCTGGGCGAGTGCATGTGCGTGTCTCTCTGTCGGGCCAGTGAAAGGTGGGTGGTTCGA	949
QY	826	CGTGCACCGAGTGTTCCTGTGACCGCCAGCTGTGGCGCGCGCTGGGAGAAACGAGGCC	885
Db	950	CGTGCACCGAGTGTTCCTGTGACCGCCAGCTGTGGCGCGCGCTGGGAGAAACGAGGCC	1009
QY	886	GGAGTGCACCAACCCCGCGCCTCTCAAGCGGGCGCTTTCTGTGAGGGGCAAGATGTC	945
Db	1010	GGAGTGCACCAACCCCGCGCCTCTCAAGCGGGCGCTTTCTGTGAGGGGCAAGATGTC	1069
QY	946	ATGACCGCACCGTCTCTCTGTGTGTCTGTGTGACGGCAGCTGGAGCCCGTGGAGCA	1005
Db	1070	AGAA---AACAGCCTGCCACACCCTGTGCCAGTGGACGCGAGCTGGAGCCCGTGGAGCA	1126
QY	1006	AGTGGTCGGCTGTGGGCTGCACTGACACCACTGGCGGAGCGGTGAGTGTCTGACCCAG	1065
Db	1127	AGTGGTCGGCTGTGGGCTGCACTGACACCACTGGCGGAGCGGTGAGTGTCTGACCCAG	1186
QY	1066	CACCCCGCAACGAGGGGAGAGTGGCAGGGCACTGACCTGGACACCCGCAACTGTACCA	1125
Db	1187	CACCCCGCAACGAGGGGAGAGTGGCAGGGCACTGACCTGGACACCCGCAACTGTACCA	1246
QY	1126	GTGACCTCTGTGTACACA-----	1143
Db	1247	GTGACCTCTGTGTACACAACCTCTACACCCCTGCCCCACCAAGGCCATGCTGTCTCCCG	1306
QY	1144	---GTGCTTCTGGCCCTGAGGAGTGCGCCCTCTATGTGGGCTCATGGCGTGGCGCTCT	1200
Db	1307	CAGCTGTCTTGGCCCTGAGGAGTGCGCCCTCTATGTGGGCTCATGGCGTGGCGCTCT	1366
QY	1201	GCCTGTCTCTGTCTGCTTGTCTCATCTCTCGTTTATTGCGGAGAGGAGGGGCTGG	1260
Db	1367	GCCTGTCTCTGTCTGCTTGTCTCATCTCTCGTTTATTGCGGAGAGGAGGGGCTGG	1426
QY	1261	ACTCAGATGTGGTGAATCGTCCATTCTCACTTCAGGCTTCAGCCCGCTCAGCATCAAGC	1320
Db	1427	ACTCAGATGTGGTGAATCGTCCATTCTCACTTCAGGCTTCAGCCCGCTCAGCATCAAGC	1486
QY	1321	CCAGCAAGCAGACACCCCCATCTGCTCAACCATCAGCCGGAACCTCAG---CACCAACA	1377

```
Db 1487 CCAGCAAGAGAGACAAACCCCATCTGCTCACCATCCAGCCGAGACCTCAGACCCACCACA 1546
Qy 1378 CCACCTACAGGAGAGTCTCTGTCCCGGAGGATGGGCCCAGGCCCAAGTTCCAGTCA 1437
Db 1547 CCACCTACAGGAGAGTCTCTGTCCCGGAGGATGGGCCCAGGCCCAAGTTCCAGTCA 1606
Qy 1438 CCAATGGGACCTGTCTAGCCCTCTGGTGGGGCGGCACACACTGACACACAGCTCTC 1497
Db 1607 CCAATGGGACCTGTCTAGCCCTCTGGTGGGGCGGCACACACTGACACACAGCTCTC 1666
Qy 1498 CCACCTCTGAGCCGAGGAGTTCTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCC 1557
Db 1667 CCACCTCTGAGCCGAGGAGTTCTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCC 1726
Qy 1558 TGCCCCGAGGACCCAGCAACATGACCTATGGACCTTCAACTTCTCTGGGGCGCGGCTGA 1617
Db 1727 TGCCCCGAGGACCCAGCAACATGACCTATGGACCTTCAACTTCTCTGGGGCGCGGCTGA 1786
Qy 1618 TGATCCCTAATACAGGTATCAGCTCTCATCCCCCAGATGCCATACCCCGAGGGAAGA 1677
Db 1787 TGATCCCTAATACAGGAATCAGCTCTCATCCCCCAGATGCCATACCCCGAGGGAAGA 1846
Qy 1678 TCTATGAGATCTACCTCAGCTGCACAAGCCGGAAGACGTGAGTTGCCCCCTAGCTGGCT 1737
Db 1847 TCTATGAGATCTACCTCAGCTGCACAAGCCGGAAGACGTG----- 1887
Qy 1738 GTCAGACCTGTGAGTCCCATCGTTAGCTGTGGACCCCTCGCGTCTGCTCACCCGCG 1797
Db 1888 -----AGCTGTGGACCCCTCGCGTCTGCTCACCCGCG 1921
Qy 1798 CAGTCATCTGCTATGAGACCACTGTGGGAGCCAGCCCTGACACCTGGAGCTCGGCC 1857
Db 1922 CAGTCATCTGCTATGAGACCACTGTGGGAGCCAGCCCTGACACCTGGAGCTCGGCC 1981
Qy 1858 TCAAAAGACAGTCTGCGAGGGCAGCTGGGAGCAGATGTGTCACCTGGGCGAGGAGS 1917
Db 1982 TCAAAAGACAGTCTGCGAGGGCAGCTGGG--AGATGTGCTGCACCTGGGCGAGGAG 2038
Qy 1918 CGCCCTCCCACTCTACTGCGAGCTGAGGCGCAGTGCCTGCTAGCTTTCACCGAGC 1977
Db 2039 CGCCCTCCCACTCTACTGCGAGCTGAGGCGCAGTGCCTGCTAGCTTTCACCGAGC 2098
Qy 1978 AGCTGGGCGCTTGGCTGGGAGAGCCCTCAGCGTGGCTGCCGCCAAGCGCTCA 2037
Db 2099 AGCTGGGCGCTTGGCTGGGAGAGCCCTCAGCGTGGCTGCCGCCAAGCGCTCA 2158
Qy 2038 AGCTGCTTCTGTGTCGCCGCTGCTGCACTCCCTCGAGTACAAATCCGGGTCTACT 2097
Db 2159 AGCTGCTTCTGTGTCGCCGCTGCTGCACTCCCTCGAGTACAAATCCGGGTCTACT 2218
Qy 2098 GCTGCAATGACACCCAGTGCATCAAGAGGTGTGAGCTGGAGAGCAGCTGGGG 2157
Db 2219 GCTGCAATGACACCCAGTGCATCAAGAGGTGTGAGCTGGAGAGCAGCTGGGG 2278
Qy 2158 GACAGCTGATCCAGGAGCCAGGCTCTGACCTTCAAGGACAGTTACCAACTCGGCC 2217
Db 2279 GACAGCTGATCCAGGAGCCAGGCTCTGACCTTCAAGGACAGTTACCAACTCGGCC 2338
Qy 2218 TATCCATCCAGATGTGCCAGCTCCCTGTGGAGAGTAAAGCTCTTGTGAGCTACAGG 2277
Db 2339 TATCCATCCAGATGTGCCAGCTCCCTGTGGAGAGTAAAGCTCTTGTGAGCTACAGG 2398
Qy 2278 AGATCCCTTTTATCAATCTGGAATGGCAGCGGTAAGTCACTGCACTGACCTTACCC 2337
Db 2399 AGATCCCTTTTATCAATCTGGAATGGCAGCGGTAAGTCACTGCACTGACCTTACCC 2458
Qy 2338 TGGAGCTGTACAGCCAGCACTAGTACCTGGCTGCAAGCTGTGGGTGGCAGGTGG 2397
Db 2459 TGGAGCTGTACAGCCAGCACTAGTACCTGGCTGCAAGCTGTGGGTGGCAGGTGG 2518
Qy 2398 AGGCGAGCGGCGAGAGCTTACGATCAACTTCAATCAATCAAGGACACAAAGGTTGCTG 2457
Db 2519 AGGCGAGCGGCGAGAGCTTACGATCAACTTCAATCAATCAAGGACACAAAGGTTGCTG 2578
```

---

```
Qy 2458 AGCTGCTGCTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGGCCCCAGTGCCTTCA 2517
Db 2579 AGCTGCTGCTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGGCCCCAGTGCCTTCA 2638
Qy 2518 AGATCCCTTCTCATTCGGCAGAGAATAATTTCCAGCTGGACCCACCTGTAGGCGGG 2577
Db 2639 AGATCCCTTCTCATTCGGCAGAGAATAATTTCCAGCTGGACCCACCTGTAGGCGGG 2698
Qy 2578 GTGCCGACTGGCGGACTGTGCCCGAGAACTCCACTGAGACCCATCTCAGCTTCTTTG 2637
Db 2699 GTGCCGACTGGCGGACTGTGCCCGAGAACTCCACTGAGACCCATCTCAGCTTCTTTG 2758
Qy 2638 CCTCAAGCCAGCCCGCAGCATGATCCTCAACTGTGGGAGCGCGGCACTTCCCCA 2697
Db 2759 CCTCAAGCCCGCAGCCCGCAGCATGATCCTCAACTGTGGGAGCGCGGCACTTCCCCA 2818
Qy 2698 ACGGCAACCTCAGCAGCTGGCTGAGCAGTGGCTGAGCTGGCCAGCAGCGTGGCC 2757
Db 2819 ACGGCAACCTCAGCAGCTGGCTGAGCAGTGGCTGAGCTGGCCAGCAGCGTGGCC 2878
Qy 2758 TCTTACAGTGTCCGAGGCTGAGTCTGA 2786
Db 2879 TCTTACAGTGTCCGAGGCTGAGTCTGA 2907
```

RESULT 6  
AAV52940  
ID AAV52940 standard; cDNA; 3014 BP.  
XX  
AC AAV52940;  
XX  
DT 25-MAR-2003 (revised)  
DT 21-DEC-1998 (first entry)  
XX  
XX Rat UNC-5 homologue unc5h-1 cDNA.  
XX  
XX UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;  
XX diagnosis; therapy; ds.  
XX  
XX Rattus sp.  
XX  
XX Key Location/Qualifiers  
XX CDS 1..2697  
XX /\*tag= a  
XX  
XX WO9837085-A1.  
XX  
XX 27-AUG-1998.  
XX  
XX 19-FEB-1998; 98WO-US003143.  
XX  
XX 19-FEB-1997; 97US-00808982.  
XX  
XX (REGC ) UNIV CALIFORNIA.  
XX  
XX Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;  
XX WPI; 1998-495364/42.  
XX P-PSDB; AAW78898.  
XX  
XX Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and  
XX the biopharmaceutical industry.  
XX  
XX Claim 7; Page 15-17; 32pp; English.  
XX  
XX This cDNA, termed unc5h-1, comprises a rat homologue of Caenorhabditis  
XX elegans unc-5. Rat unc5h-1 and unc5h-2 (see AAV52942) cDNAs were isolated  
XX from an E18 brain cDNA library. The predicted proteins (see AAW78998 and  
XX AAW78900) show similarity with UNC-5. They are predicted to be involved  
XX in cell migration and axon guidance, and are characterised as receptor  
XX proteins for netrins. Gene expression is observed in regions where  
XX differentiating neurons are undergoing axogenesis. Human unc5h-1 (see

AAV52941) and unc5H-2 (see AAV52943) cDNAs are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells by utilising these vertebrate UNC-5 nucleic acids. The invention also provides unc-5 hybridisation probes and primers, vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression) and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.). (Updated on 25-MAR-2003 to correct PI field.)

SQ Sequence 3014 BP; 596 A; 977 C; 849 G; 592 T; 0 U; 0 Other;

Query Match	Score	DB 2	Length
77.6%	2235.8	DB 2	3014

Best Local Similarity 88.5%; Pred. No. 0;  
Matches 2474; Conservative 0; Mismatches 212; T-adj-2 = 12

Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4:

QY	87	ATGCGCCTCGCGCCCGCCCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTGTGGCTC	146
Db	1	ATGCGCCTCGCGCCCGCCCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTGTGGCTC	60
QY	147	CGGGCTCGGGTGGCCAGCAGAGTGCCACCGTGGCCAAACCGTGGCTGTGGTGCACACCG	206
Db	61	CGTGTTCGGGTGGCCAGCAGAGTGCCACCGTGGCCAAACCGTGGCTGTGGTGCACACCG	120
QY	207	GACCTGTTCCCCACTTCTCTGGTGGAGCCCGAGGATGTGTACATCGTCAAGAACAGGCCA	266
Db	121	GACCTGTCGCCACATCTCTGTAGAGCTTGAGGACGTGTACATGTCAAGAACAGGCCG	180
QY	267	GTCTCTGTGTGTGAAGGCCGTGCCCGCCACGCGATCTTCTTCAAGTGCACAGGGGAG	326
Db	181	GTGTTGTGTGTGCAAGGCTGTCTCTGCCACCCAGATCTTCTCAAGTGCATGGGAA	240
QY	327	TGGGTGGCCAGTGGACACACGATGTCGCGCGAGCAGACAGACGGGACGATGTGTAGCCG	386
Db	241	TGGGTGGCCAGTGCATCAGTAAATTGAACGAGCAGACCGACAGCAGCGGATGTGCCA	300
QY	387	ACCATGGAGTCCGCAATTAAATGTCTCAAGGCAGCAGGTTCGAGAAGTGTTCGGCTGGAG	446
Db	301	ACCATGGAGTCCGTATCAACGTATCGAGGCAGCGTAGAGAAAGTGTTCGGCTGGAG	360
QY	447	GAATACTGGTGCAGTGCCTGGCATGAGTCTCTCGGCGACCAACGAAGTGCAGAGGCC	506
Db	361	GAATACTGGTGCAGTGTGGCATGAGTCTCTCGGTCACCAACGAAGTGCAGAGGCC	420
QY	507	TACATCCGCATAGCCAGATTGGCAAGAACTTCGACGAGGAGCGCTGGCCAGGAGGTG	566
Db	421	TACATCCGATTGCCATTTCGCAAGAACTTTGACGAGGAGCCATGGCCAGGAAAGTG	480
QY	567	TCCCTGGACAGGGCATCGTGTGCCCTCGCTCCGCTCACCGGAGGCACTCCCTCCAGCCGAG	626
Db	481	TCACCTGGACAGGCAATTGTACTACCTTGTGCGCCGCCAGAGGAATCCCCCGAGCTGAG	540
QY	627	GTGGAGTGGCTCCGAAACGAGACCTGTGTGGACCGCTCCCTGGACCCCAATGTATCATC	686
Db	541	GTGGAGTGGCTTCGAAATGAGGACCTGTGTGGACCCCTCCCTCGATCCCAATGTATCATC	600
QY	687	ACGCGGAGCACAGCCTGTGTGGCAGCAGGCGCGCTTGTGTGACAGCGCCAACTACACC	746
Db	601	ACGCGGAGCACAGCCTAGTGTGGTCAAGCCCGCTGGCCGACACAGCGCAACTACACC	660
QY	747	TGCGTGGCCAAAGAACATCGTGGCAAGTGTGGACCGGCTCGCTGTGTATCGTCTAC	806
Db	661	TGTGTGGCCAAAGAACATCGTAGCCGCTCGCCGAAGCACCTCTGTGACGGGTCTATTTAT	720
QY	807	GTGAACGGTGGTGTGTGACGTGGACCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	866
Db	721	GTGAACGGTGGTGTGTGACGTGGACCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	780
QY	867	GGCTGGCAGAAACGAGGCGGAGCTGCACAAACCGCGCGCTCTCAACGGGGCGGCTTTC	926

```
QY 2004 GAGGCCCTCAGCGTGGCTGCGGCCAAGCGCCTCAAGCTGCTTCTTTGGCGCGTGGCC 2063
DB 1915 GAGGCCCTCAGCGTGGCTGCGGCCAAGCGCCTCAGCTCTTCTTTGGCTCGCGTGGCC 1974
QY 2064 TGCACCTCCCTCAGTACAACTCGGCTCTACTGCTGCATGACACCCACGACCTC 2123
DB 1975 TGTACGTCCTTGTAGTACAACTCGAGTGTACTGCCTACAGACACCCACGACCTCTC 2034
QY 2124 AAGGAGGTGGTGCAGCTGAGAGAGAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTC 2183
DB 2035 AAGGAGGTGGTGCAGCTGAGAGAGAGCTAGTGGACAGCTGATCCAGGAGCCTCGGTC 2094
QY 2184 CTGCACCTTCAAGGACAGTTACCAACCTGCGGCTATCCATCCAGATGTGCCAGCTCC 2243
DB 2095 CTGCACCTTCAAGGACAGTTACCAACCTTACCTTGGAGGCGATCAACGCCAGCACCG 2154
QY 2244 CTGTGAGAGTAAAGCTCTTCTGCTACGTACAGGAGATCCCTTTTATCACATCTGGAAT 2303
DB 2155 CTGTGAGAGTAAAGCTCTTCTGCTACGTACAGGAGATCCCTTTTACCATCTGGAAC 2214
QY 2304 GGCAGCGCAGCGTACTTGCACCTTGCACCTTGCAGCGGTGTGACGCCCGACCTAGT 2363
DB 2215 GGCAGCGCAGCGTACTTGCACCTTGCACCTTGCAGCGGTGTGACGCCCGACCG 2274
QY 2364 GACCTGGCTGCAAGCTGTGGTGTGGAGGTGGAGGGGAGCGGAGAGCTTCAGCATC 2423
DB 2275 GACCTGGCTGCAAGCTGTGGTGTGGAGGTGGAGGGAGTGGGAGAGCTTCAACATC 2334
QY 2424 AACTTCAACATCACCAGGACACAAAGTTTGTGAGCTGTGCTGTGAGAGTGAAGCG 2483
DB 2335 AACTTCAACATCACCAGGACACAAAGTTTGTGAGCTGTGCTGTGAGAGTGAAGCG 2394
QY 2484 GGGGTCCCGACCGCTGTGGGCCCGCCAGTGCCTTCAAGATCCCTTCTCATTTGGCGAGAAG 2543
DB 2395 GGGGTCCCGACCGCTGTGGGCCCGCCAGTGCCTTCAAGATCCCTTCTCATTTGGCGAGAAG 2454
QY 2544 ATAATTTCAGCTGACGCCACCTGTAGCGGGGTGGCGAGTGGCGGACTTGGCGCCAG 2603
DB 2455 ATCATCGCCAGTGTGGACCCACCTGTCAGCGGGGCGCGACTTGGAGAACTTAGCCCGAG 2514
QY 2604 AACTCCACTGACAGGACATCTCAGCTTCTTTGCTTCCAAAGTCCCTTCCAGCCCGCCAGCCATG 2663
DB 2515 AACTTTCAGCTGACGCCACCTTAGCTTCTTTGCTTCCAAAGTCCCTTCCAGCCCGCCAGCCATG 2574
QY 2664 ATCCTCAACCTGTGGAGGCGCGGCACTTCCCAACCGGCAACCTTCAGCCAGCTGGGTGCA 2723
DB 2575 ATCCTCAACCTATGGGAGGCAAGGCACTTCCCAACCGGCAACCTTCAGCCAGCTGGGAGCA 2634
QY 2724 GCAGTGGCTGGACTGGGCGAGCCAGAGCTGGGCTCTTCAAGTGTGGAGGCTGAGTGC 2783
DB 2635 GCTGTGGCGGACTGGGCGCAACAGATGTGGGCTCTTCAAGTGTGGAGGCGGAGTGT 2694
QY 2784 TGAGGCGGCGCAGGCGCG--ACACCTACACTCTCACCAGCTTGGCAGCCACCAAGGACA 2841
DB 2695 TGAGACAGCGGCGGCGGTAATGGCTACATTTCTACAGCTTTGACACCTTGCAGGAGCA 2754
QY 2842 GGCAGAGCGGAGACAGGCGGCTTCCCGCACACCGGGG 2878
DB 2755 GGCAGAGCGGAGACAGGCGGCTTCCCGCACACCGGGG 2791
```

## RESULT 7

```
AA516843
ID AA516843 standard; cDNA; 2697 BP.
XX
AC AA516843;
XX
DT 14-FEB-2002 (first entry)
XX
DE Rat netrin receptor UNC5H1 (YSG7) cDNA.
XX
KW YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;
```

```
KW local cerebral glucose utilisation; phosphodiesterase 1-alpha; UNC5H1;
KW calcium-independent alpha-latrotoxin receptor; CIRL; trkB; synapsin 1A;
KW epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic; ss;
KW tumour necrosis factor alpha; TNF-alpha; rat.
XX
OS Rattus sp.
XX
Key Location/Qualifiers
CDS 1..2697
FT /*tag= a
FT /product= "Rat netrin receptor UNC5H1"
XX
EN WO200175440-A2.
XX
PD 11-OCT-2001.
XX
PF 02-APR-2001; 2001WO-GB001486.
XX
PR 31-MAR-2000; 2000GB-00007880.
PR 26-MAY-2000; 2000GB-00012768.
XX
XX (WELF-) WELFIDE CORP.
PA Cochran S, Paterson G, Ohashi Y, Morris B, Pratt J;
PI WPI: 2002-010813/01.
XX P-PSDB; AAU10543.
XX
PT Novel chronic animal model of schizophrenia, useful for identifying anti-
PS psychotic drugs and genes that are associated with schizophrenia.
XX
XX Claim 1; Fig 8a; 79pp; English.
XX
CC The invention relates to YSG polynucleotide fragments for use in
CC diagnosing and/or developing treatments for schizophrenia using chronic
CC animal models. The polynucleotides and their encoded polypeptides are
CC used for identification of compounds which modulate the expression of YSG
CC molecules, leading to the manufacture of schizophrenia medications. The
CC sequences can also be used for testing candidate compounds for any effect
CC on the polypeptides. Anti-schizophrenic effects of a compound can be
CC determined by measuring local cerebral glucose utilisation (LCGU) or
CC comparing its expression level with that of a control group. The
CC sequences are useful in the identification of genes associated with
CC schizophrenic states and in the development of an antibody. The sequences
CC of the invention include phosphodiesterase 1-alpha, calcium-independent
CC alpha-latrotoxin receptors (CIRL)-1,243, epithelial discoidin domain
CC receptor 1 (trkB), netrin receptor (UNC5H1), synapsins 1A and 1B and
CC tumour necrosis factor (TNF) alpha. This sequence represents rat netrin
CC receptor UNC5H1 (YSG7) DNA
XX
SQ Sequence 2697 BP; 541 A; 864 C; 766 G; 526 T; 0 U; 0 Other;
XX
Query Match 76.0%; Score 2189.4; DB 6; Length 2697;
Best Local Similarity 88.9%; Pred. No. 0;
Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;
QY 87 ATGCGCGTCCGCGCCCGCGCTGTGGCAGCGCTCTCTGGGATAGTCTCTCGCGCTTGGCTC 146
DB 1 ATGCGCGTCCGCGCCCGCGCTGTGGCAGCGCTCTCTGGCAGTGTCTCTCGCGCTTGGCTC 60
QY 147 CGCGGCTCGGGTGCAGCAGAGTGCACCGTGGCCAAACCCAGTGTGTGTGCAACCCG 206
DB 61 CGTGGTTCCGGTGCAGCAGAGTGCACCGTGGCCAAATCCAGTGTGGCGGCGCAACCC 120
QY 207 GACCTGTTCCCACTTCTCTGTGGAGCGCGGAGTGTACATCGTCAAGAACAGCA 266
DB 121 GACCTGTTCCCACTTCTCTGTGGAGCGCTGTGGAGCGTGTACATTTGCAAGAACAGCCG 180
QY 267 GTGCTGTTGTGTGCAAGGCGGTGCCCGCCACGAGATCTTTTCAAGTGAACCGGAG 326
DB 181 GTGTTGTGTGTGCAAGGCGGTGCCCGCCACCGAGATCTTTTCAAGTGAACCGGAA 240
QY 327 TGGGTGCGCCAGGTGGACCAAGTGTGATCGAGCGGAGCAGACAGCGGAGCAGTGGTGGCCG 386
```











Query Match		41.4%; Score 1193.6; DB 4; Length 1321;
Best Local Similarity		97.8%; Pred. No. 1.4e-210;
Matches 1295; Conservative		0; Mismatches 19; Indels 10; Gaps 8;
QY	1476	CACACACTGACACAGCTCTCCACCTCTGAGCGCGAGGTTCTCTCCGCTCTCC 1535
Db	1	CACACACTGACACAGCTCTCCACCTCTGAGCGCGAGGTTCTCTCCGCTCTCC 60
QY	1536	ACCAGAACTACTTCCGCTCCCTGCCCGAGGACACAGCAATGACCTATGGACCTTC 1595
Db	61	ACCAGAACTACTTCCGCTCCCTGCCCGAGGACACAGCAATGACCTATGGACCTTC 120
QY	1596	AATCTCTCGGGCGCGCTGATGATCCCTAATACAGGTATCAGCTCTCTCAATCCGCCA 1655
Db	121	AATCTCTCGGGCGCGCTGATGATCCCTAATACAGGTATCAGCTCTCTCAATCCGCCA 180
QY	1656	GATGCCATACCCGAGGAGATCTATGATCTACTCAGCTGCACAAAGCGGAGAC 1715
Db	181	GATGCCATACCCGAGGAGATCTATGATCTACTCAGCTGCACAAAGCGGAGAC 240
QY	1716	GTGAGTTGCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTTAGCTGTGACCC 1775
Db	241	GTGAGTTGCCCTAGCTGGCTGTGAGCCCTGCTGAGTCCCATCGTTAGCTGTGACCC 300
QY	1776	CCT-GGCGTCTCTCTCAACCGGCCAGTCTATCCT-GGCTATGACCACTGT-GGGGAGCCC 1832
Db	301	CCTGGCGTCTCTCTCAACCGGCCAGTCTATCCTGGGTATGACCACTGTGGGAGCCC 360
QY	1833	AGCCCTGACAGCT-GGAGCTTGGCTCAAAAGCAGTCTGTGGAGGAGCTGGAGCA 1891
Db	361	AGCCCTGACAGCTGGAGCTTGGCTCAAAAGCAGTCTGTGGAGGAGCTGGG--A 417
QY	1892	GGATGTCTGCACTGGGCGAGGAGCGCCCTCCACCTCTACTCTGCTCCAGCTGGAGC 1951
Db	418	GGATGTCTGCACTGGGCGAGGAGCGCCCTCCACCTCTACTCTGCTCCAGCTGGAGC 477
QY	1952	CAGTGCCTGCTACGTCTTACCGAGCAGCTGGGCGCTTTGGCCCTGGTGGGAGGCGCT 2011
Db	478	CAGTGCCTGCTACGTCTTACCGAGCAGCTGAGCGCTATGCCCTGGTGGGAGGCGCT 537
QY	2012	CAGCGTGGTGGCGCAAGCGCTCAAGTGTCTTCTGTTTGGCGCGTGGCTGACCTTC 2071
Db	538	CAGCGTGGTGGCGCAAGCGCTCAAGTGTCTTCTGTTTGGCGCGTGGCTGACCTTC 597
QY	2072	CCTCAGTACAACTCCGGCTCTACTGCTGATGACACACCGAGTGCATCAAGAGGT 2131
Db	598	CCTCAGTACAACTCTGCTCTACTGCTGATGACACTCACGATGACTCAACGTAGT 657
QY	2132	GGTGAGTGGAGAGCAGCTGGGGGACAGCTGATCCAGGACCGCGGCTCTGCATTT 2191
Db	658	GGTGAGTGGAGAGCAGCTGACGGACAGCTGATCCAGGAGCCACTGGTACTGCATTT 717
QY	2192	CAAGGACAGTTACCAACACTGGCGCTATCCATCCAGATGTCGCGAGCTCCCTGTGAA 2251
Db	718	CAAGGACAGTTACCAACACTGGCGCTATCCATCCAGATGTCGCGAGCTCCCTGTGAA 777
QY	2252	GAGTAAGCTCTTGTGCTACAGAGATCCCTTTTATCACATCTGGAATGGCACCA 2311
Db	778	GAGTAAGCTCTTGTGCTACAGAGATCCCTTTTATCACATCTGGAATGGCACCA 837
QY	2312	GCGTACTTGCATGCACTTCAACCTCGAGCGTGTACGCCCGAGCACTAGTACTGGC 2371
Db	838	GCGTACTTGCATGCACTTCAACCTCGAGCGTGTACGCCCGAGCACTAGTACTGGC 897
QY	2372	CTGCAAGCTGGTGTGGAGTGGAGGCAAGGAGGAGTTCAGATCAACTTCAA 2431
Db	898	CTGCAAGCTGGTGTGGAGTGGAGGAGGAGTTCAGATCAACTTCAA 957
QY	2432	CATCAAGGACACAAAGTTTGTGAGTCTGCTGCTGAGAGTGAAGCGGGTCC 2491
Db	958	CATCAAGGACACAAAGTTTGTGAGTCTGCTGCTGAGAGTGAAGCGGGTCC 1017
QY	2492	AGCCCTGGTGGGCCAGTGGCTTCAAGATCCCTTCTCATTCCGGCAGAGATAATTTC 2551
Db		

Db	1018	AGCCCTGGTGGGCCCCAGTGGCTTCAAGATCCCTTCCATTCGGCAGAGATAATTTC 1077
QY	2552	CAGCCTGGACCCACCCCTGTAGGCGGGTCCGACTGGCGGACTCTGGCCAGAACTCCA 2611
Db	1078	CAGCCTGGACCCACCCCTGTAGGCGGGTCCGACTGGCGGACTCTGGCCAGAACTCCA 1137
QY	2612	CTTGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCACAGCCATGATCTCAA 2671
Db	1138	CTTGACAGCCATCTCAGCTTCTTTGCTTCCAAAGCCAGCCACAGCCATGATCTCAA 1197
QY	2672	CTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGAGTGGC 2731
Db	1198	CTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGAGTGGC 1257
QY	2732	T-GGACTGGGCCAGCAGACGCTGGCCTC-TTCAAGTGT-TGGAGGCTGAGTGTCTGAGG 2788
Db	1258	TGGACTGGGCCAGCAGAGCGTGGCTTCTTTCACAGTGTTCGGAGGCTGAGTGTCTGAGG 1317
QY	2789	CCGG 2792
Db	1318	CCGG 1321

RESULT 11  
AAS75738  
ID AAS75738 standard; cDNA; 3646 BP.  
XX  
AC AAS75738;  
XX  
AC AAS75738;  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #11542.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG11551.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 11542; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders











Db 1185 TGGCATCTCATGCGGTGGGGTGGTGTACCGCGCACTGCGGTGACTTCGACA 1244  
QY 1264 CAGATGTGGCTGACTCGTCCATT---CTCACTCAGGCTTCAGCCCGTACGATCAAGC 1320  
Db 1245 CAGACATCACTGACTCATCTGCTGCCCTGACTGGTGGTTTCCACCCCGTCAACTTTAAGA 1304  
QY 1321 CCAGCAAGCAGACAACCCCTCTGCT-----CACCATCAGCGGACCTCAGACCA 1374  
Db 1305 CGCAAGGCCAGTAACCCCGCAGCTCTACACCCCTCTGTGCTCTCTGACCTGACAGCCA 1364  
QY 1375 CCACACACTACAGGCGACTCTGTTCGCCGACGATGGGCCAGCCCAAGTTCCAGC 1434  
Db 1365 GCGCGGATCTACCGCGGACCCGTGTATGCCCTGACGACTCCACCGACAAATCCCCA 1424  
QY 1435 TCACCAATGGGCACTGCTCAGCCCC-----CTGGGTGGCGCGCC 1460  
Db 1425 TGACCAACTCTCTCTGTCTGGACCCCTTACCCAGCCTTAAGGTCAAGGTCTACAGTCCA 1484  
QY 1461 -----CTGGGTGGCGCGCC 1476  
Db 1485 GCACCAAGGCTCTGGGCCAGGCTGGGAGATGGGGCTGACTGCTGGGGGTCTTGGCGC 1544  
QY 1477 ACACACTGCACACAGCTCTCCACCTCTGAGCGGAGGAGTTGCTCTCCGCTCTCCA 1536  
Db 1545 CTGGCACATACCTAGCGATTTCGCCCGGACACCACTCTCTGCACTGCGCAGCGCA 1604  
QY 1537 CCCGAACTAC-----TTCCGCTCCCTGCCCGAGGACACGAAATGACTTATGGGA 1590  
Db 1605 GCCTGGTTCCAGCAGCTCTTGGGCTGCCCGAGACCCAGGAGCAGCTCAGCGGCA 1664  
QY 1591 CTTTCAACTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCC 1650  
Db 1665 CTTTGGCTGCTCTGGGTGGAGGCTCAGCATCCCGGACAGAGGTCAGCTTGTGGTGC 1724  
QY 1651 CCCAGATGCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCGCAAGCCGG 1710  
Db 1725 CCAATGAGGCATTCCTCCAGGCAAGTTCTACAGATGATCTACTCATCAAGGCAG 1784  
QY 1711 AAGACGTGAGGTTCCTCTAGCTGTGCTGTGACACCTGTGATGCCATCGTTAGTGTG 1770  
Db 1785 AAGTACCTTCGCGCTTTTCCAGAGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844  
QY 1771 GACCCCTGCGCTCTGCTCACCGGCGAGTATCCTGCTATGACCACTGTGGGGAGC 1830  
Db 1845 GACCCACAGCCTCTCTGTGCGCGCCGCTATCTCTACCATGCCCTCTGCGCAG 1904  
QY 1831 CCAGCCTGACACTGGAGCTCGGCTCCTCAAAAGCAGTCTGTGCGAGGCGAGCTGGGAGC 1890  
Db 1905 TCAGTCCCTGACTGGATCTTTCAGCTCAAGACCCAGGCCACAGGGCCACTGGGAG- 1963  
QY 1891 AGGATGTGCTGCACTGGGGAGGAGCGGCTTCCCACTCTACTACTGCGACTGGAGG 1950  
Db 1964 --CAGGTGTGTGACCTCTGGATGAGGAGACCTGAACACACCCCTCTACTGCGACTGGAGC 2021  
QY 1951 CAGTGGCTCTAGCTTTACCGACAGCTGGGCGGCTTTCCTGCTGGGAGAGGCC 2010  
Db 2022 CCAGGGCTGTACATCTCTGCTGGACAGCTGGGCACTTACGTGTTACCGGGGAGTCT 2081  
QY 2011 TCAGCGTGGCTGCGGCCCAAGCGCTCAAGCTGCTTCTGTTGGCGCGGTGGCTGCACT 2070  
Db 2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGTGGCCGCTTTCGCCCGCCCTCTGCACCT 2141  
QY 2071 CCTTCAGTACAACATCCGGGTCTATGCTGCTGATGACACCCAGATGACTCAAGGAG 2130  
Db 2142 CCTTGGAGTACAGCTCCGGGTCTACTGCTGGAGGACACGCTGTAGCACTGAAGGAGG 2201  
QY 2131 TGTGTCAGCTGGAGAACAGCTGGGGGACAGCTGATCCAGGAGCCACGGGTCTCTGCACT 2190  
Db 2202 TGTGAGCTGGAGCGGACTCTGGCGGATCTTGTGGAGGAGCGCAACCCGCTAATGT 2261  
QY 2191 TCAAGGACAGTTACCAACCTGGCCTATCCATCCAGATGTGCCAGCTCCTGTGGA 2250

Db 2262 TCAAGGACAGTTACCAACCTGCGCTCTCCCTCCATGACCTCCCCCATGCCCATTTGA 2321  
QY 2251 AAGTAACTCTCTGCTAGCTACAGGAGATCCCTTTTATCACAATCTGGAAATGCGACGC 2310  
Db 2322 GGAGCAAGCTGCTGGCCAAATACAGGAGATCCCTTTCTATCACAATTTGGAGTGGCAGCC 2381  
QY 2311 AGCGTACTTTCGACTGCACCTTTCACCTTGGAGCGTGTGAGCCCAAGCACTAGTGAACCTGG 2370  
Db 2382 AGAAGCCCTCCACTGCACTTTCACCTTGGAGGACACAGCTTGGCTTCCACAGAGCTCA 2441  
QY 2371 CCGTCAAGCTGTGGTGTGGAGGTGGAGGGGAGCGGAGAGCTTTCAGCATCACTTCA 2430  
Db 2442 CCGTCAAGATCTGCGTGGCGCAAGTGAAGGGGAGGGCCAGATATTCAGCTTGCATACA 2501  
QY 2431 ACATCACCAG--GACACAAGTTTCTGAGCTCTGCTGCTGAGAGTGAAGCGGGG 2487  
Db 2502 CTCTGGCAGAGACACTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACTG 2561  
QY 2488 TCCCAAGCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCCCTTCTGCTGCTGGCAAGATAA 2547  
Db 2562 TCACCACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGCTCCATCCGCCAGAAATAT 2621  
QY 2548 TTTCCAGCTGGAACCCACCTGTAGCGGGTGGCGAGCTGCGCGGACTCTGGCCCAAGAAC 2607  
Db 2622 GCAACAGCTAGATGCCCCCACTCACGGGCAATGACTGGCGGATTTAGCACAGAGC 2681  
QY 2608 TCCACTGACAGCACTCTCAGCTTCTTTGCTTCAAGCCAGCCAGCCCAAGCATGATCC 2667  
Db 2682 TCTTATGACCGGTACTGAAATTTGCTCCACCAAGCAGAGCCCAAGGCTGTATCC 2741  
QY 2668 TCAACCTGTGGAGGCGCGGCACTTCCCAACCGCAACTCAGCCAGCTGGTGTGACGAG 2727  
Db 2742 TGGACTCTGGGAAGCTCTGACGAGGACGATGGGACCTCAACAGCTTGGCGAGTGCCT 2801  
QY 2728 TGCTGGACTGGGCCAGCAGCAGCTGGCTCTTACAGTGTGGAGGCTGAGTGTGA 2786  
Db 2802 TGGAGGAGATGGCAAGAGTGAAGTGTGGTGTGGCTGCTGCGCCACCGAGGACTGCTGA 2860

## RESULT 14

ABK92062  
ID ABK92062 standard; DNA; 2995 BP.

AC ABK92062;

XX DT 14-AUG-2002 (first entry)

XX DE DNA encoding novel UNC5 receptor-like protein.

XX KW Cancer; metabolic disorder; skin disorder; infectious disease; anorexia;  
XX KW behavioral disorder; valve disease; endocrine disorder; heart disorder;  
XX KW blood disorder; anxiety disorder; brain disorder; inflammatory disorder;  
XX KW neurodegenerative disorder; Alzheimer's disease; Parkinson's disease;  
XX KW immune disorder; haematopoietic disorder; dyslipidaemia;  
XX KW metabolic disturbance; obesity; metabolic syndrome X; wasting disorder;  
XX KW gene therapy; transgenic animal; human; gene; ds.

XX OS Homo sapiens.

XX PN WO200229058-A2.

XX PD 11-APR-2002.

XX PF 05-OCT-2001; 2001WO-US031248.

XX PR 05-OCT-2000; 2000US-0238323P.

XX PR 05-OCT-2000; 2000US-0238325P.

XX PR 06-OCT-2000; 2000US-0238372P.

XX PR 06-OCT-2000; 2000US-0238373P.

XX PR 06-OCT-2000; 2000US-0238379P.

XX PR 06-OCT-2000; 2000US-0238382P.

XX PR 06-OCT-2000; 2000US-0238383P.

XX PR 06-OCT-2000; 2000US-0238384P.



Db 1401 AGCCAGCCCGGATCTACCGGGACCGCTGTATGCCCTGCAGGACTCCACCGACAAAT 1460  
Qy 1430 CCAGCTACCAATGGGCACTGCTCAGCCCC----- 1460  
Db 1461 CCCATGACCAACTCTCTCTGCTGGACCCCTTACCAGCCCTTAAGGTCAAGGTCTACAG 1520  
Qy 1461 -----CTGGGTGGCG 1471  
Db 1521 CTCACGACCAACGGGCTCTGGGCCAGGCTGCGAGATGGGGCTGACCTGCTGGGGTCTT 1580  
Qy 1472 CGGCCACACCTGACACAGCTCTCCACCTCTGAGCGGAGGAGTTCGTCTCCCGCT 1531  
Db 1581 GCGGCTGCGACATACCTAGCGATTTCCCGCGGACACCCACTCTCTGCACTGGGAG 1640  
Qy 1532 CTCACACC-----AGAACTACTTCGGTCTCCCTGCCCCGAGGCAACGAGCAATGACCTTA 1585  
Db 1641 CGCCAGCCTCGGTTCCAGCAGCTCTTGGGCTGCCCCGAGACCCAGGAGGAGCGTCTAG 1700  
Qy 1586 TGGGACCTTCACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCTCTCT 1645  
Db 1701 CGGCACTTGGCTGCTGGTGGGAGGCTCAGCATCCCCCGGACAGGTGTGAGTTGCT 1760  
Qy 1646 CATCCCCCAGATGCTATACCCCGGAGGAGATCTATGAGATCTACCTCAGCTGCAAA 1705  
Db 1761 GGTGCCAATGAGCCATTTCCCGGCGAGTTCTACGAGATGTATCTCATCAACAA 1820  
Qy 1706 GCGGAAGACGTAGGTTGCCCTAGCTGGCTGTACAGCCCTGCTGAGTCCCATGTTAG 1765  
Db 1821 GCGAAGATACCTCGCGCTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGGTGAC 1880  
Qy 1766 CTGTGACCCCTGGGCTCTGCTACCGGCGAGTCACTCTGGCTATGACCACTGTGG 1825  
Db 1881 CTGTGACCCACAGGCTCTGCTGTGCGGCGGCTCACTCCACACCTGCTGCTGCTG 1940  
Qy 1826 GGAGCCAGCCCTGACAGCTGAGGCTGCGCTCAAAAGAGCAGTCTGCGAGGCGAGCTG 1885  
Db 1941 CGAAGTCAGTCCCGTCACTGATCTTTTCACTCAGACCCAGGCCACAGGGCCACTG 2000  
Qy 1886 GGAGCAGGATGTCTGCACTCGGCGAGGAGGCGCCCTCCACCTCTACTGCGCAGCT 1945  
Db 2001 GGAGCAGGAGGTGGTGACCTGGATGAGGAGACCTTGAACACACACCTGCTACTGCGCAGCT 2060  
Qy 1946 GGAGGCGAGTCTGCTAGCTCTTACCGAGAGCTGGCGGCTTTGCCCTGGTGGAGA 2005  
Db 2061 GGAGCCAGGCGCTGCATCTCTGTGACACAGCTGGGACCCCTAGCTGTACGCGCGA 2120  
Qy 2006 GCGCCTCAGCTGGCTGCGGCCAAGCGCTCAAGCTGCTTCTGTTGCGCGGCTGGCCTG 2065  
Db 2121 GTCTATTCCGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTCTTCGCCCGCCCTCTG 2180  
Qy 2066 CACCTCCCTCGAGTACAAATCCGGGTCTACTGCTGATGACACCCAGATGCATCAA 2125  
Db 2181 CACCTCCCTGGATACAGCTCCGGTCTACTGCTGGAGGACACGCTGTAGCAGTAA 2240  
Qy 2126 GGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGACAGCTGATCAGGAGCCACGGTCT 2185  
Db 2241 GGAGGTGGTGGAGCTGGAGCGGACTCTGGCGGATATCTGGTGAGAGCGCCAAACCGCT 2300  
Qy 2186 GCATTTCAAGGACAGTTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCCCT 2245  
Db 2301 AATGTTCAAGGACAGTTACCAACCTGCGCCTCTCCCTCCATGACCTCCCATGCCCA 2360  
Qy 2246 GTGGAAGATGAGCTCTGTGAGCTACAGGAGATCCCTTTATCAGATCTGGATGG 2305  
Db 2361 TTGGAGGAGCAAGCTGTGGCCAAATACCAAGAGATCCCTTCTATCAGATTTGGAGTGG 2420  
Qy 2306 CACGCGCGTACTTGCAGTGCACCTTCACTCTGGAGCGTGTGAGCCCGAGCACTAGTGA 2365  
Db 2421 CAGCGAAGGCCCTTCCAGTGCATTTTCACTCTGGAGAGGACAGCTTGGCTCCACAGA 2480  
Qy 2366 CCTGGCCTGCAAGCTGTGGGTGTGGCAGGTGGAGGGCGACGGGAGAGCTTTCAGATCAA 2425  
Db 2481 GCTCACTCTCAAGATCTGCGTGGCGCAAGTGGAGGGAGGGCGAGATATTCAGCTGCA 2540

Qy 2426 CTTCAACATCACAGGACA---CAAGGTTTGTGAGCTGTGGCTCTGGAGAGTGAAGC 2482  
Db 2541 TACCACTTGGAGAGACACCTGTGGCTCCTTGACACTCTCTGCTCTGCCCTGGCAG 2600  
Qy 2483 GGGGTCCAGACCTGTGGGCCCGAGTGCCTTCAAGATCCCTTCCCTCAATTCGGAGAA 2542  
Db 2601 CACTGTACACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAA 2660  
Qy 2543 GATAATTTCCAGCCTGGAGCCCAACCTGTAGGGGGGTGCCGACTGGGAGCTCTGCCCA 2602  
Db 2661 GATATGCAACAGCTTAGATGCCCACTACAGGGGCAATGACTGGCGGATTTAGCACA 2720  
Qy 2603 GAACTCCACCTGGAGACGCATCTCAGCTTCTTCCCTCAAGCCAGCCGCCACAGCCAT 2662  
Db 2721 GAAGCTCTCTATGAGCCGCTACCTGAATCTTTGCCACCAAGCGGCCCCACGGTGT 2780  
Qy 2663 GATCCTCAACCTGTGGAGCGCGCACTTCCCAAGCAACCTCAGCAGCTGGCTGC 2722  
Db 2781 GATCCTGACCTCTGGAAAGCTCTGACGAGGACGATGGGACCTCAACAGCTGGCGAG 2840  
Qy 2723 AGCAGTGGTGTGACTGGGCGCAGACGCTGGCTCTTCAAGTGTGGAGGCTGAGTG 2782  
Db 2841 TGCTTGGAGGAGATGGGCAAGAGTGTGTGTGGCTGTGGCCACCGAGGGAGCTG 2900  
Qy 2783 CTGAGGC 2789  
Db 2901 CTGAGCC 2907

RESULT 15  
ABO93898  
ID ABO93898 standard; DNA; 2895 BP.  
XX ABO93898;  
XX AC  
XX 01-NOV-2002 (first entry)  
XX Human transmembrane receptor UNC5H2-like NOV11 DNA, SEQ ID NO:37.  
DE Human; NOVX; neurological disorder; Alzheimer's disease;  
KW Huntington's disease; Parkinson's disease; pain; behavioural disorder;  
KW addiction; tuberous sclerosis; cancer; immune disorder; allergy;  
KW autoimmune disease; myasthenia gravis; asthma; arthritis; diabetes;  
KW thyroiditis; cardiovascular disease; hypertension; reproductive disorder;  
KW endometriosis; incontinence; psoriasis; scleroderma; alopecia; ulcer;  
KW pancreatitis; cirrhosis; glomerular endotheliosis; bacterial infection;  
KW polycystic kidney disease; endocrine disorder; obesity; cardiomyopathy;  
KW atherosclerosis; cell signal processing-related disorder;  
KW metabolic pathway regulation disorder; cytostatic; neuroprotective;  
KW antiinflammatory; immunosuppressive; analgesic; antithrombotic;  
KW dermatological; antibacterial; antiarthritic; hepatotropic; neurogenesis;  
KW differentiation; proliferation; motility; haematopoiesis; wound healing;  
KW angiogenesis; forensic biology; transgenic animal; drug screening;  
KW gene therapy; NOV11; transmembrane receptor UNC5H2-like; chromosome 10;  
KW gene; ds.  
XX Homo sapiens.  
OS

Key Location/Qualifiers  
5'UTR 1..30  
FT /\*tag= a  
FT 31..2868  
FT /\*tag= b  
FT /product= "NOV11"  
FT 31..108  
FT /\*tag= c  
FT 109..2865  
FT /\*tag= d  
FT /product= "Mature NOV11 protein"  
FT 2869..2895  
FT /\*tag= e  
XX

9.

QY 1228 TCCTCGTTTATTCGCGAAGAGGAGGGCTGAGCTCAGATGTGGTGACTCGTCAATT- 1286  
DB 1217 TGGTGGTGTACGCGCGCACTCCCGTGAATTTCGACACAGACATCACTGACTCATCTGCTG 1276  
QY 1287 --CTCACCTCAGGCTTCCAGCCCGTCCAGCATCAAGCCAGCAAAAGCAGCAACCCCGCATC 1344  
DB 1277 CCCTGACTGGTGGTTTCCACCCCGTCACTTTAAGAGCGCAAGGCCAGTAACCCGCGAGC 1336  
QY 1345 TGCT-----CACCATCCAGCCGAGCCTCAGCACACACACACCTACAGGGGAGTCTCT 1398  
DB 1337 TCCTACACCCCTCTGTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1396  
QY 1399 GTCCCGGAGAGTGGGCGCCAGCCCAAGTTCAGCTCACCACCAATGGGACCTGCTCAGCC 1458  
DB 1397 TGTATGCTGAGGACTCCACGCAAAATCCCATGACCAACTCTCTCTGCTGGACC 1456  
QY 1459 CC----- 1460  
DB 1457 CTTTACCAGCCTTAAGGTCAAGGTCTACAGCTCCAGCACCAGGGCTCTGGGCCAGGCC 1516  
QY 1461 -----CTGGGTGGCGCGGCCACACACTGACACAGCTCTCCCA 1500  
DB 1517 TGGCAGATGGGCTGACCTGCTGGGGTCTTTCGCGCTGGCAGATACCTTAGCGAFTTCG 1576  
QY 1501 CCTCTAGAGCCGAGGAGTTCGCTCCGCTCTCCACCC-----AGAACTACTCTCCGCT 1554  
DB 1577 CCGGGACACCACTTCTCTGCACTCGCAGCGCCAGCTCTGGTTCGAGAGCTCTTGG 1636  
QY 1555 CCTGCTCCGAGGACCAAGCAATGACCTATGGGACCTTCAACTTCTTGGGGGCGCGC 1614  
DB 1637 GCTGCCCCGAGACCCAGGGAGCAGGTCAGCGGCACCTTTTGGCTGCTGGTGGGAGGC 1696  
QY 1615 TGATGATCCCTAATACAGGTATCAGCTCTCTCTATCCCGCAGATGCGATACCCCGAGGGA 1674  
DB 1697 TCAGCATCCCGGCACAGGTGTGAGCTTGTGCTGGCCCAATGAGAGCAATCCCGCAGGGCA 1756  
QY 1675 AGATCTATGAGTCTACCTCACGCTGCACAAGCCGGAAGAGTGTGAGTTGCCCTAGCTG 1734  
DB 1757 AGTTCTACGAGATGTATCTACTCATCAACAGGACAGAGTACCTTGCCGCTTTCAGAG 1816  
QY 1735 GCTGTCCAGACCTCTGCTGAGTCCCATCGTTAGCTGTGGACCCCTCGGCGTCTCTCACCC 1794  
DB 1817 GGCACACAGACAGTATTGAGCCCTCGGTGACCTGTGGACCCACAGGCTCTCTGCTGTC 1876  
QY 1795 GGCAGTCACTCTGGGTTGGAACAAGTGGGAGCCAGCCCTGACAGTGTGAGCTGACCTGTC 1854  
DB 1877 GCGCCGCTCATCTCAACATCCCGCACTGTGCCAAAGTCACTGTCGCGGAGTCACTGATCTTC 1936  
QY 1855 GCCTCAAAAGACAGTGTGGAGGAGCTGGGAGCAGGATGTGCTGCACTGGGCGAGG 1914  
DB 1937 AGTCTAAGACCCAGGCCACAGGCCACTTGGG---AGGAGTGGTGACCCCTGGATGAGG 1993  
QY 1915 AGGCGCCCTCCACCTCTACTACTGCGAGCTGGAGGCCAGTGTGCTGCTGCTCTTACCG 1974  
DB 1994 AGACCTGTAACACACCTCTGCTGCTGAGCCAGGCGCTGTGACATCTGCTGTCG 2053  
QY 1975 AGCAGTGGGCGGCTTGGCCCTGTGGAGAGGCCCTCAGGTTGGCTGCGCGCCAGGCC 2034  
DB 2054 ACCAGTGGGCACTACGTGTTACGGGCGAGTCTTATTCGCTCAGCAGTCAAGCGGC 2113  
QY 2035 TCAAGCTGCTTCTGTTGGCCGCTGGCTGCACTCCCTCGAGTCAAAATCCGGGTCT 2094  
DB 2114 TCAGCTGGCGGCTTTCGCCCCCGCTCTGCACTTCCCTGGAGTACAGCTCCGGGTCT 2173  
QY 2095 ACTGCTGTGATGACACCCAGATGCACTAAGGAGGTGGTGCACTGGAGAGCAGTGG 2154  
DB 2174 ACTGCTGGAGGACACGCTGTAGCACTGAAGGAGGTGCTGGAGCTGGAGCGGACTCTGG 2233  
QY 2155 GGGGACAGTGTATCCAGGAGCCAGGGTCTGCACTTCAAGGACAGTTACCACAACTGC 2214  
DB 2234 GCGGATATCTTGGTGGAGGCGGAAACCGCTAATGTTCAAGGACAGTTACCACAACTGC 2293  
QY 2215 GCCTATCCATCCAGATGTGCCAGCTCCCTGTGGAAGAGTAAAGTCTCTTGTGAGTACC 2274

DB 2294 GCCTCTCCTCCATGACCTCCCCCATGCCATTGGAGGAGCAAGCTGTGTGGCAAAATACC 2353  
QY 2275 AGGAGATCCCTTTTATCAGACTCTGGAATGGCAACGAGGGTACTTGCACCTGACCTTCA 2334  
DB 2354 AGGAGATCCCTTTTATCAGACTTGGAGTGGCAGCCAGAGCCCTTCCACTGCATTTCA 2413  
QY 2335 CCCTGGAGGCTGTCAAGCCCAAGCACTAGTGAACCTGCTGCAAGCTGTGGGTGTGGCAGG 2394  
DB 2414 CCCTGGAGGACAGCTTGGCTCCACAGAGCTCACCTGCAAGATCTGCGTGGGCAAG 2473  
QY 2395 TGGAGGCGAACGGGAGAGCTTTCAGATCAACTTCAATCAATCAACCAAG---GACACAGGT 2451  
DB 2474 TGGAGGGGAGGGCCAGATATTCAGCTGCATACCACTCTGGCAGAGACACCTGTGCTG 2533  
QY 2452 TTGCTGAGTGTGCTGCTGAGAGTGAAGCGGGGTCCAGCCCTGTGTGGGCCCCAGTG 2511  
DB 2534 CCCTGGACACTCTGCTGCTGCTGCTGGCAGCAGCTGTCAACCCAGCTGGGACCTTATG 2593  
QY 2512 CCTTCAAGATCCCTTTCCTTTCGCGAGAGATATTTCCAGCCTGGACCCCACTGTGA 2571  
DB 2594 CTTTCAAGATCCCACTGTCTCCATCCGCCAGAGATATGCAACAGCTTAGATGCCCAACT 2653  
QY 2572 GCGGGGTGCGGACTGCGGGACTCTGGGCCCAAGAACTCCACCTGGACAGCCATCTCAGCT 2631  
DB 2654 CACGGGGCAATGACTGGCGGATGTTAGCACAGAAAGCTCTCTATGGACCGGTACTGAAAT 2713  
QY 2632 TCTTTGCTTCCAGCCAGCCCGCCAGCCATGATCTCAACCTGTGGGAGGGCGGCACT 2691  
DB 2714 ACTTTGCCCAAAAGCGAGCCCGGCTGTGATCTCTGGACCTCTGGGAGCTCTGAGC 2773  
QY 2692 TCCCCAAACGGCAACCTCAGCCAGCTGGCTGCGAGCTGGCTGGACTGGGCGAGCCAGAG 2751  
DB 2774 AGGACGATGGGACCTTCAACAGCTGCGGAGTGCCTTGGAGGAGATGGGCAAGAGTGAGA 2833  
QY 2752 CTGGCCTCTTCAAGTGTGCGAGGCTGAGTGTGAGGC 2789  
DB 2834 TGCTGGTGTGTGGCCACCGACGGGAGTGTGCTGAGCC 2871

Search completed: October 5, 2004, 15:38:48  
Job time : 1125 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 15:11:27 ; Search time 199 Seconds  
(without alignments)  
8034.239 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctgggtccggctgag.....cttccacacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA: \*  
1: /cgn2.6/prodata/2/ina/5A COMB.seq: \*  
2: /cgn2.6/prodata/2/ina/5B COMB.seq: \*  
3: /cgn2.6/prodata/2/ina/6A COMB.seq: \*  
4: /cgn2.6/prodata/2/ina/6B COMB.seq: \*  
5: /cgn2.6/prodata/2/ina/PCUS COMB.seq: \*  
6: /cgn2.6/prodata/2/ina/backfiles1.seq: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2235.8	77.6	3014	2	US-08-808-982-1
2	2235.8	77.6	3014	3	US-09-306-902A-1
3	1570.4	54.5	1787	2	US-08-808-982-2
4	1570.4	54.5	1787	3	US-09-306-902A-2
5	1147.8	39.8	1282	4	US-09-833-381-1806
6	783.4	27.2	2831	2	US-08-808-982-3
7	783.4	27.2	2831	3	US-09-306-902A-3
8	419	14.5	1605	4	US-09-833-381-1807
9	256	8.9	771	1	US-08-253-155A-17
10	119	4.1	305	2	US-08-808-982-4
11	119	4.1	305	3	US-09-306-902A-4
12	51.8	1.8	1280	3	US-08-483-533-38
13	51.8	1.8	1280	4	US-09-283-471A-38
14	51.2	1.8	699	4	US-09-252-991A-7947
15	51.2	1.8	1302	4	US-09-252-991A-7947
16	51.2	1.8	1947	4	US-09-252-991A-7809
17	50.8	1.8	1770	4	US-09-252-991A-7533
18	50.8	1.8	3885	4	US-09-369-364A-16
19	50.8	1.8	4884	4	US-09-252-991A-12126
20	50.8	1.8	4884	4	US-09-252-991A-12126
21	50.2	1.7	1300	3	US-08-483-533-39
22	50.2	1.7	1300	4	US-09-283-471A-39
23	48.6	1.7	1068	4	US-09-252-991A-9933
24	48.6	1.7	1251	4	US-09-252-991A-10142
25	48	1.7	4257	2	US-08-690-473-1
26	48	1.7	4257	3	US-09-259-821A-1
27	48	1.7	4257	3	US-08-843-659-1

c	28	48	1.7	8312	4	US-09-620-312D-1048	Sequence 1048, Ap
	29	47.8	1.7	1288	1	US-08-440-856A-9	Sequence 9, Appli
c	30	47.8	1.7	2211	4	US-09-252-991A-11256	Sequence 11256, A
	31	47.8	1.7	2289	4	US-09-252-991A-10995	Sequence 10995, A
	32	47.6	1.7	1627	4	US-09-675-018B-11	Sequence 11, Appl
	33	47	1.6	7218	1	US-08-212-463-14	Sequence 14, Appl
	34	46.8	1.6	1951	3	US-08-222-865-1	Sequence 1, Appli
	35	46.8	1.6	1951	4	US-09-510-949-1	Sequence 1, Appli
	36	46.4	1.6	453	4	US-09-252-991A-4004	Sequence 3971, Ap
	37	46.4	1.6	567	4	US-09-252-991A-4004	Sequence 4004, Ap
	38	46.2	1.6	11220	3	US-09-105-537-32	Sequence 32, Appl
	39	46.2	1.6	36778	3	US-09-105-537-5	Sequence 5, Appli
	40	46.2	1.6	38506	3	US-09-320-878-19	Sequence 19, Appl
	41	46.2	1.6	38506	4	US-09-141-908-1	Sequence 1, Appli
	42	46.2	1.6	38506	4	US-09-657-440-19	Sequence 19, Appl
	43	46	1.6	1065	4	US-09-340-798A-31	Sequence 31, Appl
	44	46	1.6	3624	1	US-07-951-715A-6	Sequence 6, Appli
	45	46	1.6	3624	2	US-08-459-448A-6	Sequence 6, Appli

## ALIGNMENTS

### RESULT 1

US-08-808-982-1  
; Sequence 1, Application US/08080982  
; Patent No. 5939271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsay  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,982  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 3014 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-808-982-1

Query Match 77.6%; Score 2235.8; DB 2; Length 3014;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;

OY 87 ATGGCCGTCGGCCCGCTGTGGCCAGCGCTCTGGGCATAGCTCCGCGCTGGCTC 146  
|||||

Db	1	ATGGCGTTCGGCCCGCGCTGTGGCAGTGTCTCTTGGGCATAGTCTCTCGCGCGCTGGCTT	60
Qy	147	CGGGCTCTGGGTGCCAGCAGAGTGCACCGTGGCCAAACCCAGTGTCTTGGTGCAACCG	206
Db	61	CGTGGTTCGGTGTGCCAGAGAGTGCACCGTGGCCAAATCCAGTGTCCGGTGCACACCC	120
Qy	207	GACCTGTTTCCCACTTCTCTGTGTGGAGCCCGAGATGTACATCTCAAGAACAGCA	266
Db	121	GACCTGTGCCCATCTTCTGGTAGAGCTTGAGGCGTGACATTTCAAGTCAAGAACAGCG	180
Qy	267	GTGCTGTTTGTGTGCAAGGCGTGCCTCCGACCGCAGATCTTCTTCAAGTGCAACGGGAG	326
Db	181	GTGTTGTGTGTGCAAGGCTGTGCCTGCCACCCAGATCTTCTTCAAGTCAATGGGAA	240
Qy	327	TGGTGGCCAGGTGTGGACACAGTGATCGAGCGACAGACAGCGGAGCAGTGGTGAGCG	386
Db	241	TGGTTCGCCAGGTTCGATCACTGAATTTGAACAGCAGCACGACAGCAGCGATTTGCCA	300
Qy	387	ACCATGGAGTCCGCATTAATGTCTCAAGCAGCAGGTTCAGAGGTGTTCGGCTGGAG	446
Db	301	ACCATGGAGTCCGTATCAAGTATCGAGCGCAGGTAGAGAAAGTGTTCGGCTGGAG	360
Qy	447	GAATACTGTGTCCAGTGTGGCATGAGCTCTCTGGGCAACCAAGAGTCAAGAGCC	506
Db	361	GAATACTGTGCCAGTGTGGCATGAGCTCTCTGGGTACCAACCAAGTCAAGAGCC	420
Qy	507	TACATCCGCATAGCAGATTTCGCAAGAACTTCGAGCAGGAGCGGTGGCCAAAGAGTG	566
Db	421	TACATCCGGAATGGCTATTTTCGCAAGAACTTTGAGCAGGAGCCATGGCCAAAGAGTG	480
Qy	567	TCCCTGGAGCAGGCGATCGTCTGCCCTCGCGTCACCGAGGCGATTCCTCAGCGGAG	626
Db	481	TCACTGGAGCAAGCATTTGATACCTTGTGCGCCCCCAGAGGAATCCCGCCAGCTGAG	540
Qy	627	GTGAGTGGCTCCGAAACGAGGACCTGGTGGACCGCTCCCTGGNCCCAATGTATACATC	686
Db	541	GTGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTGTACATC	600
Qy	687	ACGGGGAGCAGCGTGGTGGTGCAGCAGGCCCGCGCTTGTGACAGGGCAACTACACC	746
Db	601	ACGGGGAGCAGCGCTAGTCTGGGTGAGCCCGCGCTGGCCGACAGGCCAACTACACC	660
Qy	747	TGCTGGCCAAAGACATCGTGGCAAGTGCCTGGAGCGGCTCCGCTGTCTGTCTGTCTAC	806
Db	661	TGTGTGGCCAAAGACATCGTAGCCCGTGCAGAGCAGCCTCTGAGCGGTCAATGTATTAT	720
Qy	807	GTGAACGTGGTGTGTGACCTGTGACCGAGTGTCCGTCTGACGCGCAGCTGTGGCGC	866
Db	721	GTGAACGTGGTGTGTGACCTGTGACCTGTGAGTGTGTCCGTCTGACGCGCAGCTGTGGCGT	780
Qy	867	GGCTGGCGAAACGAGCGCGAGCTGCACCAACCGCGCGCTCTCAACGGGGGCGCTTTC	926
Db	781	GGCTGGCGAAACGAGCGCGAGCTGCACCAACCGCGCACCTCTCAACGGGGGCGCGCTTC	840
Qy	927	TGTGAGGGCGAAGTCCCATGACCGCACCGTCTCTCTCTGTGTGTCTGTGTGAGCGC	986
Db	841	TGTGAGGGCGAAGTCCGAAACAGCGCTGGCGCACTCTG---TGCACCATGGATGGG	897
Qy	987	AGCTGGAGCCGTGGAGCAAGTGTGTGGCTGTGGCTGGACTGCACCACTGGCGGAGC	1046
Db	898	AGCTGGAGTTCGTGGAGTAAAGTGTGTGAGCTGTGGCTGTGACTGCACCACTGGCGGAGC	957
Qy	1047	CGTGAAGTGTGTGACCCAGCACCCCGAAACGAGGGGAGAGTGCAGGCGACTGAGCTG	1106
Db	958	CGCGAGTGTCTGTGACCCAGCACCCCGCAATGGAGGTGAGGAGTGTGGGGTGTGACCTG	1017
Qy	1107	GACACCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCGCTTGGAGACGTG	1166
Db	1018	GACACCGCAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCGCTTGGAGACGTG	1077
Qy	1167	GCCTCTATGTGGGCGCTCATATCGCGTGTGCGCTGTGCTGTCTGTCTGTCTGTCTGCTC	1226
Db	1078	GCTCTCTATACGGCGCTGTGCTGTGGCTGTGCGCTCTTCTTGTCTGTGTGGCGCTT	1137







QY 2788 GCGGCGCAGGCGCCGACACCTACCTCTCACAGCTTTGGCA--CCCAAGGACAGCA 2845  
DB 1672 GCGGCGCAGG--CGAACACTACATTTTACAGTTTGGGAACCCCAAGGACAGCA 1729  
QY 2846 GAAGCGGACAGGCGCCCTTCCACACACGCGGAG 2881  
DB 1730 GAAGCGGACAGGCGCTTTTCCAAAACCGGGAGA 1765

## RESULT 4

US-09-306-902A-2

; Sequence 2, Application US/09306902A

; Patent No. 6277585

; GENERAL INFORMATION:

; APPLICANT: Tessier-Lavigne, Marc

; Leonardo, E. David

; Hink, Lindsay

; Masu, Masayuki

; Kazuko, Keino-Masu

; TITLE OF INVENTION: Netrin Receptors

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESS: SCIENCE &amp; TECHNOLOGY LAW GROUP

; STREET: 268 BUSH STREET, SUITE 3200

; CITY: SAN FRANCISCO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94104

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/306,902A

; FILING DATE: 07-May-1999

; CLASSIFICATION: &lt;Unknown&gt;

; ATTORNEY/AGENT INFORMATION:

; NAME: OSMAN, RICHARD A

; REGISTRATION NUMBER: 36,627

; REFERENCE/DOCKET NUMBER: UC96-217

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 343-4341

; TELEFAX: (415) 343-4342

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1787 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA

; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 54.5%; Score 1570.4; DB 3; Length 1787;  
Best Local Similarity 96.9%; Pred. No. 0;  
Matches 1721; Conservative 0; Mismatches 36; Indels 19; Gaps 11;

QY 1114 GCAACTGTACCACTCTGTGTACACAGTCTTCTGGCCCTTGAGGAGCTGGCCCTCT 1173  
DB 1 GCNACTGTACCACTGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 59  
QY 1174 ATGTGGGCTCTATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1233  
DB 60 ATGTGGGCTCTATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 119  
QY 1234 TTTATTGCCGAAGAGAGGAGGCGCTGGACTCAGATGGTGTGCTGTGCTGTGCTGTGCTGT 1293  
DB 120 TTTATTGCCGAAGAGAGGAGGCGCTGGACTCAGATGGTGTGCTGTGCTGTGCTGTGCTGT 179  
QY 1294 CAGGCTTCCAGCCGCTCAGCATC-AAGCCGAGCAAGCAGCAACCCCATCTGCTCACC 1352

DB 180 CAGGCTTCCAGCCGCTCAGCATCTAAGCCCAAGCAGCAACCCCATCTGCTCACC 239  
QY 1353 ATCCAGCCGAGCCTCAG---CACCACACCACTTACAGGGCAGTCTCTGTCCCCGCGAG 1409  
DB 240 ATCCAGCCGAGCCTCAGCACCACCACTTACAGGGCAGTCTCTGTCCCCGCGAG 299  
QY 1410 GATGGGCCAGCCCCCAAGTTTCAGCTCACCATATGGGCACTTCTCAGCCCCCTTGGTGGC 1469  
DB 300 GATGGGCCAGCCCCCAAGTTTCAGCTCACCATATGGGCACTTCTCAGCCCCCTTGGTGGC 359  
QY 1470 GGCGCCACACACTGACACAGCTCTCCACCTCTGAGGCGGAGAGTTCGTCTCCCGC 1529  
DB 360 GGCGCCACACACTGACACAGCTCTCCACCTCTGAGGCGGAGAGTTCGTCTCCCGC 419  
QY 1530 CTCTCCACCCAGAACTACTTCGGCTCCCTGCCCGCAGGCAACCAATGACCTATGGG 1589  
DB 420 CTCTCCACCCAGAACTACTTCGGCTCCCTGCCCGCAGGCAACCAATGACCTATGGG 479  
QY 1590 ACCTTCAACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCATC 1649  
DB 480 ACCTTCAACTTCTCTCGGGGCGGCTGATGATCCCTAATACAGGAATCAGCCTCTCATC 539  
QY 1650 CCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACTCAGCTGCAACAGCG 1709  
DB 540 CCCCAGATGCCATACCCCGAGGAGATCTATGAGATCTACTCAGCTGCAACAGCG 599  
QY 1710 GAAGACGTGAGTTGGCCCTAGCTGGCTGCTGAGACCTCTGAGTCCCATCGTTAGCTGT 1769  
DB 600 GAAGACGTGAGTTGGCCCTAGCTGGCTGCTGAGACCTCTGAGTCCCATCGTTAGCTGT 659  
QY 1770 GGACCCCTTGGCTCTGCTCACCGGCGAGTCATCTCGGCTATGACCACTGTGGGAG 1829  
DB 660 GGACCCCTTGGCTCTGCTCACCGGCGAGTCATCTCGGCTATGACCACTGTGGGAG 719  
QY 1830 CCCCAGCTGACAGCTGGAGCTGGCCCTCAAAAAGCAGTCTGCGAGGCGAGCTGGGAG 1889  
DB 720 CCCCAGCTGACAGCTGGAGCTGGCCCTCAAAAAGCAGTCTGCGAGGAG-AGCTGGGAG 778  
QY 1890 CAGGATGTGCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG 1949  
DB 779 GATG---TCTGCACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCCAGCTGGAG 834  
QY 1950 GCGAGTGCCTGCTACGTTCTTCCAGCAGCTGGGCGGCTTTGCCCTGTGGGAGAGGCC 2009  
DB 835 GCGAGTGCCTGCTACGTTCTTCCAGCAGCTGGGCGGCTTTGCCCTGTGGGAGAGGCC 894  
QY 2010 CTCAGCTGCTGCGCCCAAGCGCTCAAGCTGCTTCTGTTGCGCGGCTGGCCTGCACC 2069  
DB 895 CTCAGCTGCTGCGCCCAAGCGCTCAAGCTGCTTCTGTTGCGCGGCTGGCCTGCACC 954  
QY 2070 TCCCTCGAGTACAAATCCGGGTCTACTGCCTGCATGACACCCAGCATGCACCTCAAGGAG 2129  
DB 955 TCCCTCGAGTACAAATCCGGGTCTACTGCCTGCATGACACCCAGCATGCACCTCAAGGAG 1014  
QY 2130 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCAGCGGTCTGCGAC 2189  
DB 1015 GTGGTGCAGCTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCAGCGGTCTGCGAC 1074  
QY 2190 TTCAAGAGAGATTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCCCTGTGG 2249  
DB 1075 TT-AAGGACAGTTTACCACACCT--GCCCTATCATCCAGATGTGCCAGCTCCCTGTGG 1131  
QY 2250 AAGAGTAAGCTCTCTGTGCTACAGGAGATCCCTTTTATCACAATCTGGAATGGCAGC 2309  
DB 1132 AAGAGTAAGCTCTCTGTGCTACAGGAGATCCCTTTTATCACAATCTGGAATGGCAGC 1191  
QY 2310 CAGCGTACTGTGCACTGACCTTTCACCTGGAGCGGTGAGCCCGAGCAGTGTGAGCTG 2369  
DB 1192 CAGCGTACTGTGCACTGACCTTTCACCTGGAGCGGTGAGCCCGAGCAGTGTGAGCTG 1251  
QY 2370 GCCTGCAAGCTGTGGTGTGGCAGGTGGAGGCGAGCGGGCAGAGCTTCAAGATCAACTTC 2429  
DB 1252 GCCTGCAAGCTGTGGTGTGGCAGGTGGAGGCGAGCGGGCAGAGCTTCAAGATCAACTTC 1311



```
QY 2430 AACATCACCAAGGACACAAAGTTTGTGAGTCTGTGGCTCTGGAGAGTGAAGCGGGGTC 2489
Db 1312 AACATCACCAAGGACACAAAGTTTGTGAGTCTGTGGCTCTGGAGAGTGAAGCGGGGTC 1371
QY 2490 CGAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATAAAT 2549
Db 1372 CCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATAAAT 1431
QY 2550 TCCAGCTGAGACCAACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCGAGAAACTC 2609
Db 1432 TCCAGCTGAGACCAACCTGTAGCGGGGTGCCGACTGGCGGACTCTGGCCGAGAAACTC 1491
QY 2610 CACCTGGACAGCCATCTCAGCTTCTTTCCTTCAAGCCAGCCACAGCATGATCCCTC 2669
Db 1492 CACCTGGACAGCCATCTCAGCTTCTTTCCTTCAAGCCAGCCACAGCATGATCCCTC 1551
QY 2670 AACCTGTGGAGGCGGGCACTTCCCAACGGCAACTCAGCCAGCTGGCTGACAGTGTG 2729
Db 1552 AACCTGTGGAGGCGGGCACTTCCCAACGGCAACTCAGCCAGCTGGCTGACAGTGTG 1611
QY 2730 GTTGGACTGGGCGAGCCAGCAGCTGGCTTTCACAGTGTGGAGGCTGATGCTGAG 2787
Db 1612 GTTGGACTGGGCGAGCCAGCAGCTGGCTTTCACAGTGTGGAGGCTGATGCTGAG 1671
QY 2788 GCCGGCCAGGCGCCGACCTACACTCTCAGCAGCTTGGCA--CCACCAAGGACAGGCA 2845
Db 1672 GCCGGCCAGG--CGAACACTACAAATTTTACAGTGTGGAGGCTGATGCTGAG 1729
QY 2846 GAAAGCGGACAGGGGCGCTTCCCAACCGGGGAGA 2881
Db 1730 GAAAGCGGACAGGGGCTTTTCCCAAAACCGGGGAGA 1765

RESULT 5
US-09-833-381-1806
; Sequence 1806, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1806
; LENGTH: 1282
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(1282)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1806

Query Match 39.8%; Score 1147.8; DB 4; Length 1282;
Best Local Similarity 97.7%; Pred. No. 3.3e-240;
Matches 1259; Conservative 0; Mismatches 18; Indels 11; Gaps 9;

QY 1510 CCGAGGAGTTGGTCTCCCGCTCTCCACCCAGAACTACTCCGCTCCCTGGCGGCGGCTGATGATCCCTAATA 1629
Db 1 CCGAGGAGTTGGTCTCCCGCTCTCCACCCAGAACTACTCCGCTCCCTGGCGGCGGCTGATGATCCCTAATA 60
QY 1570 CCAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATA 1629
Db 61 CCAGCAACATGACCTATGGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATA 120
QY 1630 CAGGTATCAGCTCTCTATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCT 1689
Db 121 CAGGAATCAGCTCTCTATCCCCCAGATGCCATACCCGAGGGAAGATCTATGAGATCT 180
```

```
QY 1690 ACCTCAGCTGCACAAAGCCGGAAGACGTGAGGTTGCCCCCTAGCTGGCTGTTCAGACCCCTGC 1749
Db 181 ACCTCAGCTGCACAAAGCCGGAAGACGTGAGGTTGCCCCCTAGCTGGCTGTTCAGACCCCTGC 240
QY 1750 TGAGTCCCATCGTTAGCTGTGGACCCCT--GGCGTCTGCTCACCCTGCGGCGGCTGATCTCT- 1807
Db 241 TGAGTCCCATCGTTAGCTGTGGACCCCTGCGGCGTCTGCTCACCCTGCGGCGGCTGATCTCTG 300
QY 1808 GGCTATGAGCCACTGTGGGAGCCCAAGCCCTGACAGCT--GGAGCTTGGCGCTCAAAAAGC 1866
Db 301 GGCTATGAGCCACTGTGGGAGCCCAAGCCCTGACAGCTGGAGCTTGGCGCTCAAAAAGC 360
QY 1867 AGTCGTGCAGGCGAGCTGGAGAGAGTGTGCTGACCTGGGCGGAGGAGGCGCCCTCCC 1926
Db 361 AGTCGTGCAGGCGAGCTGGAGAGAGTGTGCT--TGCACCTGGGCGAGGAGGCGCCCTCCC 418
QY 1927 ACCTCTACTACTGCCAGCTGGAGGCGAGTGCCTGCTACGTTCTTACCGAGCAGCTGGGCC 1986
Db 419 ACCTCTACTACTGCCAGCTGGAGGCGAGTGCCTGCTACGTTCTTACCGAGCAGCTGAGCC 478
QY 1987 GCTTTGCCCTGGTGGAGAGGCGCTCAGCGTGGCTGCCCGCAAGCGCCTCAAGCTGCTTC 2046
Db 479 GCTATGCCCTGGTGGAGAGGCGCTCAGCGTGGCTGCCCGCAAGCGCCTCAAGCTGCTTC 538
QY 2047 TGTTTGGCGCGTGGCTGCTGACCTCCCTCGAGTAAACATCCGGGTCTACTGCTGCATG 2106
Db 539 TGTTTGGCGCGTGGCTGCTGACCTCCCTCGAGTAAACATCTGGTCTACTGCTGCATG 598
QY 2107 ACACCCAGATGCACCTCAAGGAGTGTGCTGAGCTGGAGAGCAGCTGGGGGAGACAGCTGA 2166
Db 599 ACACCTCAGATGCACCTCAAGCTGAGTGTGCTGAGCTGGAGAGCAGCTGGGGGAGACAGCTGA 658
QY 2167 TCCAGGAGCCAGGCTCCTGCACTTCAAGGACAGTTACCAACACCTGCCCTTATCCATCC 2226
Db 659 TCCAGGAGCCAGGCTCCTGCACTTCAAGGACAGTTACCAACACCTGCCCTTATCCATCC 718
QY 2227 AGGATGCGCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAAGTACCAAGGAGATCCCT 2286
Db 719 AGGATGCGCCAGCTCCCTGTGGAAGAGTAAGCTCCTTGTCAAGTACCAAGGAGATCCCT 778
QY 2287 TTTATCACATCTGGAATGGCAGCGGTAATTGCACTGCACCTTCACTTGGAGCGTGG 2346
Db 779 TTTATCACATCTGGAATGGCAGCGGTAATTGCACTGCACCTTCACTTGGAGCGTGG 838
QY 2347 TCAGCCCGCAGCACTAGTGACCTGGCTTCAAGCTGTGGTGTGGCAGTGGAGGCGAGC 2406
Db 839 TCAGCCCGCAGCACTAGTGACCTGGCTTCAAGCTGTGGTGTGGCAGTGGAGGCGAGC 898
QY 2407 GCGAGAGCTTCAAGCATCACTTCAACATCAACAGGACACAAAGGTTTGTGAGCTGCTGG 2466
Db 899 GCGAGAGCTTCAAGCATCACTTCAACATCAACAGGACACAAAGGTTTGTGAGCTGCTGG 958
QY 2467 CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCT 2526
Db 959 CTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAAGATCCCT 1018
QY 2527 TCCTCATTTGGCAGAGATAA--TTTCCAGCTTGGACCCACCTGTAGCGGGGTGCCGAC 2585
Db 1019 TCCTCATTTGGCAGAGATAA--TTTCCAGCTTGGACCCACCTGTAGCGGGGTGCCGAC 1078
QY 2586 TGGCGG--ACTCTGCGCCAGAAATCTCCACCTGGAGCAGCACTCTAGCTTCTTTGGCTCAA 2644
Db 1079 TGGCGGAACTCTGGCGCCAGAAATCTCCACCTGGAGCAGCACTCTAGCTTCTTTGGCTCAA 1138
QY 2645 GCCCAGCCCGCAGCAGCACTTCACTTCAAGCTTGGAGGCGGCGGCACTTCCCAAGCGGAA 2704
Db 1139 GCCCAGCCCGCAGCAGCACTTCACTTCAAGCTTGGAGGCGGCGGCACTTCCCAAGCGGAA 1198
QY 2705 CCTCAGCCAGCTGGCTGCGAGCAGTGGCTGGACTGGGCGAGCAGCAGCTGCTCTTAC 2764
Db 1199 CCTCAGCCAGCTGG--TGCAGCAGTGGCTGGACTGGGCG--AGCAGACGCTGG--CTCTTAC 1254
```

QY 2765 AGTCGAGGCTGAGTCTGAGGCCG 2792  
DB 1255 AGTGTGCGAGGCTGAGTCTGAGGCCG 1282

## RESULT 6

US-08-808-982-3  
; Sequence 3, Application US/08808982  
; Patent No. 5939271  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; APPLICANT: Leonardo, E. David  
; APPLICANT: Hink, Lindsay  
; APPLICANT: Masu, Masayuki  
; APPLICANT: Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/808,982  
; FILING DATE:  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2831 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
US-08-808-982-3

Query Match 27.2%; Score 783.4; DB 2; Length 2831;  
Best Local Similarity 59.4%; Pred. No. 5.7e-161;  
Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;

QY 184 ACCAGTGCCTGGTGCCAAACCCGAGCTGCTCCCACTTCCTGTTGGAGCCCGAGGATG 243  
DB 104 ACTCCTTCCCATCAGCACCGCGGAGCAGCTGCTCACTTCCTGCTGGAACCAAGAGATG 163  
QY 244 TGTACATCGTCAAGAACAGCCAGTGTGTTGTGTGTCAGAGCCGTTGCCCGCCACGAGA 303  
DB 164 CCTACATCGTTAAAGAACAGCCAGTGTGAATTCACCTGCCAGAGCTTCCTTCGCCACACAGA 223  
QY 304 TCTTCTTCAGTGCACAGGGGAGTGGTGCGCCAGGTGGACCACTGATCGAGCGAGCA 363  
DB 224 TCTACTTCAAGTGTATGGCGAGTGGTTAGCCAGAAAGGCCACGTCACCCAGAGGCC 283  
QY 364 CAGACGGGAGCAGTGTGTAGCCGACCATGGAGTTCGCAATTAATGTCTCAAGGCAGCAGG 423  
DB 284 TGGATGAGGCCACAGGCTTCGCAATACGAGAGTGCAGATAGAGGTGTGCGCGCAGCAGG 343  
QY 424 TCGAAGAGGTGTTCGGGCTGGAGGAATACTGTGTGCCAGTGGTGGCATGGAGTCTCCTCGG 483  
DB 344 TGGAGGAACACTTTTGGGCTCGAGGACTACTGGGTGTACGTGCGTGGCTGTGGAGTCTTTCGG 403

QY 484 GCACCACCAAGAGTCAGAAAGCCCTACATCCCATATAGCCAGATTGGCGAAGAACTTCGAGC 543  
DB 404 GAAACCAAGAGTCGCGAGCCTACATCCGCTTACCTGCGCAAGAACTTTGACC 463  
QY 544 AGGAGCCGCTGGCCAAAGAGGTGTCCTTGGAGCAGGGCATCGTGTGCTGCCCTGCCCTCCAC 603  
DB 464 AGGAGCCTCTGGCGAAGGAGGTACCTTTGGATCATGAGGTCTCTGAGTGGCCGCCAC 523  
QY 604 CGGAGGCATCCCTCCAGCCGAGGTGGTCTCCGGAACGAGGACCTGTGTGAGCCCGT 663  
DB 524 CAGAGGAGTGCCTGTGGCTGAGGTGAATGGCTCAAGAAAGAATGTCATCGATCCCG 583  
QY 664 CCTGGACCCCAATGTATACATCACGCGGAGCACAGCCTGTGTGTGGACAGGCCGCC 723  
DB 584 CTCAGGACACTAACTTCCTGCTCACCATTTGACCAACCTCATCATCGCCAGGCGGCC 643  
QY 724 TTGCTGACAGGCCAACTACACCTCGTGGCCCAAGAACATCGTGGCATCGTCCGCCAGCG 783  
DB 644 TCTCAGACACAGCCAACTACACCTGTGTGGCAAGAAATATTGTGGCCAAAGCCCGGAGCA 703  
QY 784 CCTCCGCTGTGTCTCATGTCTACGTGAACGGTGGTGGTCCGACGTGGACCGAGTGTCCG 843  
DB 704 CGACGGCCACAGTCACTGCTCTATGTGAACGGAGGTGGTCCAGCTGGGCGAGATGGTCA 763  
QY 844 TCTGACGCGCCAGCTGTGGCGCGCTGGCAGAAACGAGCCGAGCTGCAACAACCCCG 903  
DB 764 CTTGCTCTAAACGCTGCGCGCGAGGTGGCAGAAACGTAAGGACCTGACCAACCCAG 823  
QY 904 CCGCTCTCAACGGGGCGCTTCTGTGAGGGGAGAAATGTCATGACCGCACCGTCTCT 963  
DB 824 CCCCACCTCAATGGAGGTGCTTCTCGAGGGGACAGGCTTGGCCAGAA---GACGGCTTGA 880  
QY 964 CTCGCTGTCTGTGGACGCGCAGCTGGAGCCGCTGGAGCCGTAAGTGGTGGCTGTGGGC 1023  
DB 881 CCACCGTGTGCCCGAGTGGATGAGCGTGGACTGAGTGGAGCAAGTGGTCCGCTCGAGCA 940  
QY 1024 TGGACTGACCCCACTGGCGGAGCGGTAGTGTCTTGACCCAGACCCCGCAACGAGGGG 1083  
DB 941 CAGAGTGTGCGCACTGGCGCAGCGAGTGCATGGCACCGCGCCGCCAAGAGCGAGGCC 1000  
QY 1084 AGGAGTGCAGGGCACTGACCTGGACCCCGCAACTGTACAGTACAGTACCTCTGTGTACACA 1143  
DB 1001 GTGACTGACGCGGAGCGCTACTTGACTCCAAGAACTGCACCCAGTGGGCTGTGCTGTGA 1060  
QY 1144 GTGCTTCTGGCC-----CTGAGGACGTGGCCCC 1170  
DB 1061 ATCAGAGAACTCTAAACAGCCCTTAAGCGCGCCCTGGAGCGCTGGGAGACGTGGCGC 1120  
QY 1171 TCTATGTGGGCCCTC---ATCGCGGTGGCGCTGTGCTGTGCTGTGCTGTGCTGTCTCA 1227  
DB 1121 TGTATCGGGGCTCGGTGGTGGCGCTCTTTGTGGTTCTGGCAGTTCTCATGGCTGTAGGAG 1180  
QY 1228 TCTCGTTTATTCGCGGAAGAGGAGGGGCTGGACTCAGATGTGGTGTGCTGTGCTGTCAATT- 1286  
DB 1181 TGTATCGGTGTACCGGAGAACTGCGGGGACTTTCACACGAGACATCACTGACTCTCTGTGCTG 1240  
QY 1287 --CTCACCTCAGGCTTCCAGCCCGCTCAGCATCAAGCCCGCAGCAAGAGAGCAACCCCATC 1344  
DB 1241 CCGTCACTGGTGTTCACCCCGCTCAACTTCAAGACTGCAAGGCCCGCAGCAACCCACAGC 1300  
QY 1345 TGTCT-----CACCATCCAGCCGAGCCTCAGCAACA---CCACCACCTACCAGGGGAGTC 1395  
DB 1301 TCTGTACCCCATCCGCGCCCTCCGAGCTTAACGCGCCAGTGTGTCATCTACCGCGGACCTG 1360  
QY 1396 TCTGTCTCCCGGAGGA----- 1411  
DB 1361 TGTATGCCCTGCGAGGACTCTGCCGCAAGATCCCTATGACTAATTACCCCTTCTGTGATC 1420  
QY 1412 -TGGGCCCAAGCCCAAGTTCCAGCTCAACCAATGGGCACTGTCTCAGCCC-----CC 1461  
DB 1421 CCTTGGCCAGGCTCAAGATCAAGGTCTATGACTCCAGCACCACTCGGCTCTGGGGCTGGCC 1480  
QY 1462 TGGGTGGCGCCGCCACACTGACCACTGCCACAGCTCTCCCACTCTGAGGCGGAGGAGTTCG 1521

1481 TGCTGATGAGCGAGCTGCTGGGTGCTTACCAACCGGTACATATCCAGGCGATTCT 1540  
1522 TCTCCCGCTCTCACCAAGC-----TACTTCGGCT 1554  
1541 CCGGGACACCACTTCTCTGCACTGCGAGCGCCAGCTTGGTTCCAGCAGCTCTCTGG 1600  
1555 CCTGCCCCGAGGACACAGCAACATGACCTATGGAGCTTCAACTTCTCGGGGCGGGC 1614  
1601 GCCTCCTCGAGACCCAGCAGCAGTGTAGTGGCACTTGGTTGCTGGTGGGAGGC 1660  
1615 TGATGATCCCTAATACAGGTATCAGCTCTCTCATCCCGCCAGATGCCATCCCGAGGGA 1674  
1661 TGACCAATTCCTCGGCACAGGGGTGAGCTGTGTGTACCAATGAGGCCATTCCTCCAGGCA 1720  
1675 AGATCTATGATGATCTACCTCACCTGCACAGCGGAGAGCTGAGGTGCCCCCTAGCTG 1734  
1721 AGTCTATGATCTGTATCTAGCTATCAACAAGACTGAAAGCACTCCCACTTTTCGGAAG 1780  
1735 GCTGTGAGACCTCTGCTGAGTCCCATCGTTAGCTGTGAGCCCTCGGCTCTCTCACCC 1794  
1781 GTTCCAGACAGTATTGAGCCCTCGGTGACCTGCGGGCCACGGGCTCTCTCTGTGCC 1840  
1795 GGCAGTCTCTGCTGATGGAACATGTGGGAGCCCGCCTGACAGCTGGAGCTGC 1854  
1841 GGCCTGTTCTCTCACTGTGCCCACTGTCTGAAGTCAATGCGGAGACTGGATCTTCC 1900  
1855 GCTCAAAAGAGTGTGCGGAGGAGCTGGAGCAGAGTGTGTCACCTGGGCGAGG 1914  
1901 AGCTAAGACCCAGGCCCATGAGGCCACTGGG---AGGAGTGTGACTTTGGATGAG 1957  
1915 AGCGCCCTCCCACTCTACTGCGAGCTGGAGCCAGTGTCTGCTCTTCCACCG 1974  
1958 AGACTCTGAACACCCCTCTACTGCGAGCTAGAGCTAATCTGCCATCTCTTGG 2017  
1975 AGAGTGGGCGCTTTGCGCTGTGGGAGAGCCCTCAGCGTGGCTGCGCCCAAGCGCC 2034  
2018 ACCAGTGGGTACTACGTGTTCACGGCGAGTCTTACTCCCGCTCCGAGTCAAGCGCG 2077  
2035 TCAAGTGTCTTGTGCGCGGTGGCTGACCTCCCTCGAGTACACATCCGGGTCT 2094  
2078 TCAGTACGCTATCTTGGCGCCAGCCCTCTGACCTCCCTGGAGTATAGTCTCAGGGTCT 2137  
2095 ACTGCTGATGACACCCAGTGTCACTCAAGAGGTGTGCACTGGAGAGAGCTGG 2154  
2138 ACTGTGAGGACACTCTGCGAGACTGAGAGGTCTAGAGCTGGAGAGACTCTGG 2197  
2155 GGGGACAGTGTATCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTACCAACCTGC 2214  
2198 GTGGCTACTTGGTGGAGGAGCCAGACTTTGCTCTTTAAGGACAGTTACCAACCTAC 2257  
2215 GCCTATCCATCCAGATGTCACGCTCCCTGCGAGAGTAACTCTTGTGAGTACC 2274  
2258 G-CTCTCTCTCTGACATCCCTCCATGCGCCACTGGAGGAGCAAACTACTGGCCAGTACC 2316  
2275 AGGATCCCTTTTATACATCTGGAATGGACGCGGCTTGTGCTGCTGCACTGCACTTCA 2334  
2317 AGGATTTCTTCTACCATGTGTGAGCGGAGCCAGCAAGCCCTGCACTGCACTTTCA 2376  
2335 CCCTGAGGTGTAGCCCGCAGCACTAGTGAACCTGGCTGCAAGCTGTGGTGTGGCAGG 2394  
2377 CCCTGGAGAGACATAGCTTACCTCCAGTGTACCTGTAAAGTGTCTGCTGCGCGCAGG 2436  
2395 TGGAGGGGACGGGAGAGCTTTCAGCATCACTTCAACATAC---CAAGGACAAAGT 2451  
2437 TAGAAGGGAAGGCGAGATTTTCCAGCTGCACACACGCTGGCTGAGAGCGCTGTGGCT 2496  
2452 TTGCTGAGCTGTGGCTCTGAGAGTGAAGCGGGGTCCACGCTGTGGTGGCCCGATG 2511  
2497 CCCTGATGCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2556  
2512 CTTCAAGATCCCTTCTCTATTCGGCAGAGATAAATTTTCAGCTGGACCCACCTCTGA 2571

2557 CTTCAAGATACCACTGTCTCCATCGGCCAGAGATCTGCAACAGCTGGAAGCCCCAACT 2616  
2572 GCGGGGTGCGACTGTGGGAGACTTGTGCCAGAAACTTCCACCTGGGACAGCCATCTCAGCT 2631  
2617 CACGGGCAATGACTGTGGGCTGTGTGGCAAGAAGCTCTCCATGAGCCGGTACTGAACT 2676  
2632 TCTTTCCTTCAAGCCAGCCCGCCAGCAGCATGATCTCAACCTGTGGGAGGCGGGGACT 2691  
2677 ACTTGGCCACCAAGAGCTAGTCCACAGGCGGTGATCTTAGACCTCTGGGAAGCTCGGCAGC 2736  
2692 TCCCCAACGCAACCTCAGCCAGCTGCTGACAGTGGCTGAGCTGGGCGCAGCCAGC 2751  
2737 AGATGATGGGACCTCAGCAGCTGCGCAGTGCCTTGGAGGAGATGGGCAAGTGA 2796  
2752 CTGGCCTCTTCAAGTGTGGAGGCTGAGTGTGA 2786  
2797 TGCTGTAGCCATGACCACTGATGGCGATTGCTGA 2831

RESULT 7  
US-09-306-902A-3  
; Sequence 3, Application US/09306902A  
; Patent No. 6277585  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsay  
; Masu, Masayuki  
; Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/306,902A  
; FILING DATE: 07-May-1999  
; CLASSIFICATION: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSWAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2831 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-306-902A-3

Query Match 27.2%; Score 783.4; DB 3; Length 2831;  
Best Local Similarity 59.4%; Pred No. 5.7e-161;  
Matches 1625; Conservative 0; Mismatches 971; Indels 139; Gaps 12;  
184 ACCAGTGCCTGGTGCACACCGGAGCTGTCTCCCACTTCTGCTGGAGCGCGAGGATG 243  
104 ACTCCTTCCCATCAGCAGCCCGGAGGAGCTGCTTCTTCTGCTGGAACAGGAGATG 163  
244 TGTACATCGTCAAGAACAGCCAGTGTGCTGTGTGTCAGAGGCGGTGCCCCCAGCAGA 303

164 CCTCATCTGTAAGAACCAAGCCAGTGGAAATTCACATCTGCGAGCCTTCCCTGCCACACAGA 223  
1304 TCTCTTTCAAGTGCAACGGGAGTGGGTGCGCAGGTGGACCAACGATGATCGAGCGCAGCA 363  
224 TCTACTTCAAGTGTAATGGCAGTGGTTAGCCAGAAAGGCCACGTCACGACGAGAGGCC 283  
364 CAGACGGAGCAGTGGTGAGCCGACCATGGAGGTTCGCATTAATGTCTCAAGAGCAGCAGG 423  
284 TGGATGAGGCCACAGGCTTTGGAAATACGAGAGGTGCAGATAGAGGTGTCGCGCAGCAGG 343  
424 TCGAGAAGGTGTTCGGCTGAGGAAATACTGGTGCAGTGCCTGGCATGGAGCTCCTCGG 483  
344 TCGAGAAGCTTTTGGCTCGAGGACTACTGGTGTGAGTGCCTGGCATGGAGCTCTTCGG 403  
484 GCACCAACCAAGAGTCAGAAAGGCTACATCCGATAGCCAGATGGCGCAAGAACTTCGAGC 543  
404 GAACCAACCAAGAGTCGCCAGGCTTACATCCGATTCGCTACTTGGCAAGAACTTTGACC 463  
544 AGGAGCGCTGCCAAGGAGGTTCCTGGAGCAGGGCATCTGCTGCCCTCGCTCCGCTCCAC 603  
464 AGGAGCGCTTGGCAGAGGAGGTACCTTGGATCATGAGGTCTTCTGCGAGTGCGCCAC 523  
604 CGGAGGCGATCCTCAGCGGAGGTGGAGTGGTCCGGAACGAGGACCTGGTGGACCCGT 663  
524 CAGAGGAGTGTCTGTGGCTGAGTGGAAATGGCTCAAGAAATGAAGATGTATCGATCCCG 583  
664 CCTGAGCCCAATGATATACATCAGCGGGAGACAGCCTGGTGGTGGACAGCCGCGC 723  
584 CTCAGGACACTAACTTCCTGCTCACCATTGACCAACCTCATCTCCGCAAGCGCGCC 643  
724 TTGCTGACACGGCAACTACCTGCTGGTGGCCCAAGAACATCGTGGACGCTGCCCGCAGG 783  
644 TCTCAGACACGCCAACTACCTGTGTGGCAAGAAATATTTGGCCAAAGCGCGCGAGCA 703  
784 CTTCCGCTGTCTCATCTGTCTAGTGAACGGTGGTGGTTCGACGTGGACCGAGTGTGTCG 843  
704 CGACGCCACAGTCATCTGTATGTGAACGGAGTTGGTCCAGCTGGGCGAGATGTGTAC 763  
844 TCTGACGCCCACTGTGGCGCGCTGGCGAGAAACGGAGCCGAGCTGACCAACCCGG 903  
764 CTTGCTTAACCGCTCGCGCGAGGTGGCAGAAAGCTATAGGACCTGCACCAACCCAG 823  
904 CGCTCTCAACGGGCGCTTTCTGTGAGGGCAGAAATGTCAATGACCGACCGTCTCT 963  
824 CCCCCTCAATGAGGTGCTTCTGCGAGGACAGGCTTGCCAGAA---GACGGCTTGA 880  
964 CTCTGTCTGTCTGTGAGCGCAGCTGGAGCCCGTGGAGCAAGTGGTGGCCCTGTGGCC 1023  
881 CCACCGTGTGCCAGTGGATGGAGCGTGGACTGAGTGGAGCAAGTGGTCCGCTGCAGCA 940  
1024 TGGACTGCACCACTGGCGAGCGGTGAGTGTCTGACCCAGACCCCGCAACGGAGGGG 1083  
941 CAGAGTGTGGCACTGGCGAGCGCGAGTGCATGGCAGCCCGCCCGCAAGAGCGGCC 1000  
1084 AGGAGTGCAGGCACTGACCTGACACCGCAACTGTATACAGTACCTCTGTGTACACA 1143  
1001 GTGACTGACGGGACGCTACTTGTCTCAGAAACTGCAACCGATGGCTGTGGTGTGCTGA 1060  
1144 GTGCTTTGTGGCC-----CTGAGGACGTGGCCC 1170  
1061 ATCAGAGAACTTAAGACCTTAAAGCCGCCCTGGAGCGCTGGGAGAGCGTGGCGC 1120  
1171 TCTATGTGGCCCTC---ATGCGCGTGGCGCTGCTGGTCTGCTGCTGCTCTCTCA 1227  
1121 TGTATGGGCGCTCGTGGTGGCGCTCTTTGTGGTTCCTGGCAGTTCATCGGCTGTAGGAG 1180  
1228 TCTCTGTTTATTCGCCGAAGAGAGGGCTGACTCAGATGTGGCTGAGTCTGTCGTCATT- 1286  
1181 TGAATGCTACCGGAGAACTCGCGGACCTTCGACACGGACATCACTGATCTCTCTGCTG 1240  
1287 --CTCAGCTCAGGCTTCCAGCCCGTTCAGCATCAAGCCCGCAGCAAGCAGACAAACCCCATC 1344

1241 CCTCACTGGTGGTTTCCACCCCGCTCAACTTCAAGACTGCAAGGCCAGCAACCCACAGC 1300  
1345 TGCT-----CACCATCCAGCGGACCTCAGACCA---CCACCACTTACAGGCGAGTTC 1395  
1301 TCTTGCAACCCATCCGCCCTCCGGACCTTAACGGCCAGTGTGGCATCTACCGCGGACCTG 1360  
1396 TCTGTCCCGCGCAGGA-----TACCTTCCCGT 1554  
1361 TGTATGCTCTGAGGACTCTGCCGACAAGATCCCTATGACTAATTCACCCCTTCTGGATC 1420  
1412 -TGGGCCAGAGCCCAAGTTCCAGCTCAACCAATGGGCACTGTCTCAGCCC-----CC 1461  
1421 CCTTCCCGACGCTCAAGATCAAGTCTATGACTCCAGCACCATCGCTCTGGGGTGGCC 1480  
1462 TGGGTGGGGCGCGCACACACTGCACACAGCTCTCCACCTCTGAGGGCGAGGAGTTCG 1521  
1481 TGGCTGATGGAGCCGACCTGTGGGTGTCTTACCAACCGGTATACCAAGCGGATTTCT 1540  
1522 TCTCCCGCTCTTCCACCCAGAAC-----TACTTCCGCT 1554  
1541 CCGGGACACCCACTTCTGTGCACCTGGCGAGCGCAGCCTTGGTTCCCGACCTCTCTCG 1600  
1555 CCTTGGCCGAGGACAGCAATGACCTATGGGACCTTCAACTTCTCGGGGCGCGC 1614  
1601 GCCTCCTCGAGACCCCGACGACGTGTCAAGTGGCACCTTTGGTTGCTGGGTGGAGGC 1660  
1615 TGATGATCCCTAATACAGGTATCAGCTCTCATCCCCCAGATGCATACCCCGAGGGA 1674  
1661 TGACCAATTCGCGGACAGGGTCAAGCTTGTGGTACCAATGGAGCCATTCGCCAGGGA 1720  
1675 AGATCTATGAGATCTACCTCAGCTGCACAAGCCGGAAGACGTGAGTTGCCCTAGCTG 1734  
1721 AGTCTATGACTTGTATCTACGTATCAACAAGACTGAAAGCACCTTCCACCTTTCGAAG 1780  
1735 GCTGTGACACCTGTGAGTCCCATGCTTGTAGTGTGGACCCCTGCGCTGCTCTCACCC 1794  
1781 GTTCCAGACAGTATTTAGGCCCTCGGTGACCTGCGGGGCCAGGGGCTCTCTCTGTC 1840  
1795 GGCAGTCACTCTGGTATGGACCACTGTGGGAGCCAGCCCTGACGTGAGGCTGTC 1854  
1841 GCCTGTGTCTCTCACTGTGCCCTCTGTGCTGAAGTCAATGCCGAGACTGATCTTCC 1900  
1855 GCTCAAAAAGCAGTCTGTGCGAGGCGAGTGGGAGCAGGATGTGCTGCACCTGGGCGAGG 1914  
1901 AGCTCAAGACCCAGGCCCATCAGGCCACTGGG---AGGAGTGTGTGACTTTGGATGAGG 1957  
1915 AGCGCCCTTCCACCTCTACTTACTGACGTGAGGCGCAGTGCCTGCTAGCTTCTCACCG 1974  
1958 AGACTTGAACACCCCTCTACTGCGAGCTAGAGCTAAATCCTGCCACATCCTGTGTG 2017  
1975 AGCAGCTGGCGCGCTTTGCCCTGGTGGGAGAGCCCTCAGCGTGGCTGCGCCAGCGCC 2034  
2018 ACCAGTGGGTACTAGTGTTCACGGGCGAGTCTCTACTCCCGCTCCGAGTCAAGCGGC 2077  
2035 TCAAGTGTCTTGTGTGGCGCGGTGGCTGCACTCTCCCTCGAGTACAACATCCGGGTCT 2094  
2078 TCCAGCTAGCACTTCTGCGCCCGCCAGCCCTCTGCACTCTCCCTGGAGTATAGTCTCAGGCTCT 2137  
2095 ACTGCTGTATGACACCCAGGATGCACTCAAGAGGTTGTGACGCTGGAGAGCAGCTGG 2154  
2138 ACTGTCTGAGGACACTCTCTGACGACTGAAGAGGTCTTAGAGCTGGAGAGACTCTGG 2197  
2155 GGGGACAGTGTATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTACCAACCTGC 2214  
2198 GTGGCTACTTGTGGAGGAGCCCAAGACTTTGCTCTTTAAGGACAGTTACCAACCTAC 2257  
2215 GCTATCCATCCAGATGTGCCAGTCTCCCTGTGGAGAGTAAAGTCTTGTGAGTACC 2274  
2258 G-CTCTCCTCTCCATGACATCCCCCATGCCACTGGAGGAGCAAACTACTGGCCAAAGTACC 2316  
2275 AGGAGATCCCTTTTATCAGATCTGGAAATGGCAGCGCGGTACTTTCAGTGCACCTTCA 2334  
2317 AGGAGATTCCTTCTTACCATGTGTGGAACCGGAGCCAGAAAGCCCTGCACTGCACTTTCA 2376



```
Patent No. 5691147
GENERAL INFORMATION:
APPLICANT: Gyuris, Jenő
APPLICANT: Draetta, Giulio
TITLE OF INVENTION: CDK4 Binding Proteins
NUMBER OF SEQUENCES: 95
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII(text)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/253,155A
FILING DATE: 02-JUN-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Vincent, Matthew P.
REGISTRATION NUMBER: 36,709
REFERENCE/DOCKET NUMBER: MII-028
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 771 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-253-155A-17

Query Match 8.9%; Score 256; DB 1; Length 771;
Best Local Similarity 67.2%; Pred. No. 1.3e-46;
Matches 406; Conservative 1; Mismatches 191; Indels 6; Gaps 3;

QY 1741 AGACCTGTGAGTCCATCGTTAGCTGTGGAGCCCTCGCGTCTCAGCCGCTCACCGGCCAG 1800
DB 738 AGACAGTATTGACCCCTCGGTGACCTGTGGACCCACAGCCCTCTGCTGCGGCCCG 679
QY 1801 TCATCTGTGCTATGGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGCGCTCA 1860
DB 678 TCATCTCACCATTGCCCACTGTGCCGAAGTCAGTGCCTGACTGGATCTTTCAGCTCA 619
QY 1861 AAGAGCAGTGTGCGAGGAGCTGGAGCAGATGTGTGACCTGGGCGAGGAGCGC 1920
DB 618 AGACCCAGCCCAACAGGCGCACTGGG---AGGAGGTGGTACCTGGATGAGGAGCCC 562
QY 1921 CTTCCACCTCTACTACTGCGAGCTGGAGCCAGTGTGCTGCTACGCTTTCACCGAGCAG 1980
DB 561 TGAACACACCTGCTACTGCGAGCTGGAGCCAGGCGCTGTCATCTCTGCTGGAACAGC 502
QY 1981 TGGGCGCTTTGCTGCTGTGGAGAGCCCTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2040
DB 501 TGGGCACTTACGCTGTTTACGCGGCGAGTCTCTATTCGCGCTCAGCAGTCAAGCGGCTCAGC 442
QY 2041 TGTCTTGTGTCGCGCGGTGGCTGACCTCCTCGAGTACACATCCGCGTCTACTGCTG 2100
DB 441 TGGCGGTCTTTCGCGCGCGCGCTCTGACCTCTCTGGAGTACAGACTCCGCGTCTACTGTC 382
QY 2101 TGCATGACACCCACGATGACCTCAAGAGGTGGTGTGCTGAGAGAGCAGCTGGGGGAC 2160
DB 381 TGGAGGACAGGCGCTGTAGCAGTGAAGAGGTGCTGGAGCTGGAGCGGACTCTTCGGCGGAT 322
QY 2161 AGCTGATTCAGGAGCCAGGCTCTGCACTTCAAGGACAGTTCACCAACCTCGCGCTAT 2220
DB 321 ACTTGTGGAGGAGCCGAAACCGCTAATGTTCAAGGACAGTTCACCAACCTCGC-CTCT 263

QY 2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGTAAGCTCCTTGTACCTACAGGAGA 2280
DB 262 TCCTCATGACTTCCCCATGCCATTTGGAGGAGCAAGCTGTGCCAATACAGGAGA 203
QY 2281 TCCCCTTTTATCATCTGTGAATGGACGACGCGGTACTTGGCACTGCACTTTCACCTGG 2340
DB 202 TCCCCTTCTATCATCTTGGAGTGGAGCCAGAGG--CCTCCACTGCACTTTCACCTGG 145
QY 2341 AGCG 2344
DB 144 AGAG 141

RESULT 10
US-08-808-982-4
Sequence 4, Application US/08080982
Patent No. 5939271
GENERAL INFORMATION:
APPLICANT: Tessier-Lavigne, Marc
APPLICANT: Leonardo, E. David
APPLICANT: Hink, Lindsay
APPLICANT: Masu, Masayuki
APPLICANT: Kazuko, Keino-Masu
TITLE OF INVENTION: Netrin Receptors
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/808,982
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A.
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 305 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cdna
US-08-808-982-4

Query Match 4.1%; Score 119; DB 2; Length 305;
Best Local Similarity 68.1%; Pred. No. 5.6e-17;
Matches 209; Conservative 0; Mismatches 95; Indels 3; Gaps 3;

QY 1906 TGGGCGAGGAGCGCCCTCCACCTCTACTTGCAGCTGGAGGCGAGTGTGCTGTACG 1965
DB 1 TGGATGAGGAGACCTTGAACACACACCTGCTACTG-CAGCTGGAGCCAGGGCTG-TACA 58
QY 1966 TTTTACCGAGCAGCTGGGCCCGTTTGCCTGTGGAGAGGCCCTTCAGCTGGTGGCGG 2025
DB 59 TCTGTGTGACCACTGGGCACCTACGTTTTCACGGGCGAGTCTATTTCGCGCTCAGCAG 118
QY 2026 CCAAGGCGCTCAAGCTGTCTTGTTCGCGCGTGGCCCTGCACTCCCTCGAGTACAACA 2085
DB 119 TCAAGCGGCTCCAGCTGGCGGT-TTCGCCCCCGCCCTCTGCACTCCCTCGGAGTACAGCC 177
```





Db 615 ACCCCGCGACCCCGCGACCCCGCGACCCCGCGCGGTCGCTTCTCGCCCCACGTC 674  
QY 126 ATAGTCCTCGCGCTTGGCTCGCGGCTCGGTCGCCAGCAGAGTGCACCGTGCCCAAC 185  
Db 675 CGGTGCGCCACCTGGTGTCTGGGCTCGCGCGCCCGCTCGCGCGCGGCTCGTGG 734  
QY 186 CAGTGTCTGGTGCACCCGACCTGCTTCCCACTTCTCTGTGGAGCCGAGGATGT 244  
Db 735 GCCCGGAGCGGCGCGACCGGCTCGGTTCCGCGCGCGGTCGCGAGGCGGCGGT 793

RESULT 13  
US-09-283-471A-38  
; Sequence 38, Application US/09283471A  
; Patent No. 6340673  
; GENERAL INFORMATION:  
; APPLICANT: Roizman, Bernard  
; TITLE OF INVENTION: Method For Treating Tumorigenic Diseases  
; NUMBER OF SEQUENCES: 43  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 6300 Sears Tower, 233 South Wacker Drive  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606-6402  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/283,471A  
; FILING DATE: 04-APR-1999  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/861,233  
; FILING DATE: 31-MAR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/419,853  
; FILING DATE: 11-APR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/483,533  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Zeller, James P.  
; REGISTRATION NUMBER: 28,491  
; REFERENCE/DOCKET NUMBER: 27373/32742A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312/474-6300  
; TELEFAX: 312/474-0448  
; TELEX: 25-3856  
; INFORMATION FOR SEQ ID NO: 38:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1280 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
US-09-283-471A-38

Query Match 1.8%; Score 51.8; DB 4; Length 1280;  
Best Local Similarity 51.0%; Pred. No. 0.034;  
Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 6 GGAGCTCGGCTGAGCGCTAAAGCGCTCCGCGCGGGCGCGCGCGCGCGCGCG 65  
Db 555 GCGGCGGGGAGGGGCGCGAGCGCCCGCGACCCCGCGACCCCGCGAGCCCGCG 614  
QY 66 CCGGCTTCGCCCGCGCGCGCATAGCGCTCCGCGCGCGCTGTGGCAGCGCTCTGGGC 125

Db 615 ACCCCGCGACCCCGCGACCCCGCGACCCCGCGCGGTCGCTTCTCGCCCCACGTC 674  
QY 126 ATAGTCCTCGCGCTTGGCTCGCGGCTCGGTCGCCAGCAGAGTGCACCGTGCCCAAC 185  
Db 675 CGGTGCGCCACCTGGTGTCTGGGCTCGCGCGCCCGCTCGCGCGCGGCTCGTGG 734  
QY 186 CAGTGTCTGGTGCACCCGACCTGCTTCCCACTTCTCTGTGGAGCCGAGGATGT 244  
Db 735 GCCCGGAGCGGCGCGACCGGCTCGGTTCCGCGCGCGGTCGCGAGGCGGCGGT 793

RESULT 14  
US-09-252-991A-7947  
; Sequence 7947, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 7947  
; TYPE: DNA  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-7947

Query Match 1.8%; Score 51.2; DB 4; Length 699;  
Best Local Similarity 47.5%; Pred. No. 0.036;  
Matches 152; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1876 AGGGCAGCTGGGAGCAGGATGTGTCACCTGGCGGAGGAGGCGCCCTCCACCTACT 1935  
Db 168 ACAGCTCTACGACGACGAGGTGCTCGAGGCCAAGTGTCTGTCGGGCGCATTCGGCT 227  
QY 1936 ACTGCCAGCTGGAGCCAGTGTCTGTCTTACCCAGCAGCTGGGCGCGCTTTGCC 1995  
Db 228 ACGACCTGGTGTGCGCAGCAGCAGCTTCTGTCACCACTACCTGAAGGCGGAGTGTTC 287  
QY 1996 TGTGGGAGAGGCGCTCAGCTGGTGGCGCCAGCGCTCAGCTGCTTCTGTTGGCG 2055  
Db 288 AGCCGCTGGACAAGACGAGTGTGCGAAGTGTGCGAAGTGTGCGGCGGCTGCTCAAGG 347  
QY 2056 CGGTGGCTTGCACCTCCCTCGAGTACAACTCGGGTCTACTGCTGCTGATGACACCCACG 2115  
Db 348 TGCTGCGCGGAGGACCCCGGCAACCGTACGTGATGCCCTACATGTGGGCGACCAACG 407  
QY 2116 ATGCACTCAAGGAGGTGTGTCAGCTGGAGAGCAGCTGGGGGACAGCTGATCAGAGC 2175  
Db 408 GCATCGCTACAACCTCGACAAGTGTGCGGCGGCTGCTCGGCGACGATGCGCGCTGACT 467  
QY 2176 CACGGTCTGCACTTCAAG 2195  
Db 468 CTGGGACCTGTGTTCAG 487

RESULT 15  
US-09-252-991A-7809  
; Sequence 7809, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18

Query Match 1.8%; Score 51.8; DB 4; Length 1280;  
Best Local Similarity 51.0%; Pred. No. 0.034;  
Matches 122; Conservative 0; Mismatches 117; Indels 0; Gaps 0;

QY 6 GGAGCTCGGCTGAGCGCTAAAGCGCTCCGCGCGGGCGCGCGCGCGCGCGCG 65  
Db 555 GCGGCGGGGAGGGGCGCGAGCGCCCGCGACCCCGCGACCCCGCGAGCCCGCG 614  
QY 66 CCGGCTTCGCCCGCGCGCGCATAGCGCTCCGCGCGCGCTGTGGCAGCGCTCTGGGC 125

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 7809
; LENGTH: 1302
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-7809
```

```
Query Match      1.8%; Score 51.2; DB 4; Length 1302;
Best Local Similarity 47.5%; Pred.No.0.046;
Matches 152; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QY 1876 AGGGCAGCTGGGAGCAGGATGCTGCACCTGGGCGAGGAGGCGCCCTCCACCTCTACT 1935
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1936 ACTGCCAGCTGGAGGCCAGTGCCTGCTAGCTCTTACCCGAGCAGCTGGGCGGCTTTGCCG 1995
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 428 ACGACCTGGTGTGCCCGCAGCAGCTTCTGCCCACTACCTGAAGGCCGAGGTGTTCC 487
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1996 TGGTGGGAGAGGCCCTCAGCGTGGCTGCCGCCAAGCGCCTCAAGCTGCTTCTTTTGGCG 2055
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 488 AGCCGCTGGACAAGAGCAAGCTGCCGAAGTGGAAAGAACCTCAACCCGGCCCTGCTCAAG 547
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2056 CGGTGGCCTGCACCTCCCTCGAGTACAACATCCGGCTCTACTGCCCTGCATGACCCACG 2115
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 548 TGCTGCCCGCAGAGNCCCGGCAACCGTACGTGATGCCCTACATGTGGGGCACCACG 607
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2116 ATGCACTCAAGAGGTGTTGTCAGCTGGAGAAGCAGCTGGGGGGACAGCTGATCCAGGAGC 2175
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 608 GCATCGCCTACAACTCGACAAGTCCGCGCGGTGCTCGGCGACGATGCGCGCTGGACT 667
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 2176 CACGGGTCTGCACTTCAAG 2195
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 668 CTTGGACCTGGTGTTCAG 687
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: October 5, 2004, 20:46:56  
Job time : 210 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 15:20:19 ; Search time 1321 Seconds  
(without alignments)

11056.338 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctggggtccgggtgag.....ccttccacacacgggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3340653 seqs, 2534783454 residues

Total number of hits satisfying chosen parameters: 6681306

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US05\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2881	100.0	2881	11	US-09-970-944-1
2	2676.4	92.9	2752	13	US-09-918-779-1
3	2676.4	92.9	2752	17	US-10-624-932-1
4	2402.6	83.4	3580	17	US-10-311-623-13
5	2235.8	77.6	3014	10	US-09-933-261-1
6	2235.8	77.6	3014	15	US-10-256-702-1
7	2189.4	76.0	2697	15	US-10-240-154-15
8	1570.4	54.5	1787	10	US-09-933-261-2
9	1570.4	54.5	1787	15	US-10-256-702-2
10	1193.6	41.4	1321	13	US-10-296-115-365
11	1147.8	39.8	1282	9	US-09-833-381-1806
12	889	30.9	2860	13	US-10-087-684-1
13	889	30.9	2860	13	US-10-218-779-1
14	887.4	30.8	2860	13	US-10-087-684-3

15	887.4	30.8	2860	13	US-10-218-779-3	Sequence 3, Appli
16	875.8	30.4	2995	13	US-09-972-211-55	Sequence 55, Appli
17	875.8	30.4	2995	13	US-10-096-625-55	Sequence 55, Appli
18	866.4	30.1	2895	13	US-10-037-417-37	Sequence 37, Appli
19	855.2	29.7	3485	9	US-09-816-828-18	Sequence 18, Appli
20	853.6	29.6	3884	13	US-10-147-493-145	Sequence 145, App
21	853.6	29.6	3884	13	US-10-145-127-145	Sequence 145, App
22	853.6	29.6	3884	13	US-10-160-503-145	Sequence 145, App
23	853.6	29.6	3884	13	US-10-143-118-145	Sequence 145, App
24	853.6	29.6	3884	13	US-10-144-993-145	Sequence 145, App
25	853.6	29.6	3884	13	US-10-158-787-145	Sequence 145, App
26	853.6	29.6	3884	13	US-10-140-024-145	Sequence 145, App
27	853.6	29.6	3884	13	US-10-140-808-145	Sequence 145, App
28	853.6	29.6	3884	13	US-10-152-405-145	Sequence 145, App
29	853.6	29.6	3884	13	US-10-127-852A-145	Sequence 145, App
30	853.6	29.6	3884	13	US-10-127-900A-145	Sequence 145, App
31	853.6	29.6	3884	13	US-10-128-685A-145	Sequence 145, App
32	853.6	29.6	3884	13	US-10-131-820A-145	Sequence 145, App
33	853.6	29.6	3884	13	US-10-142-886-145	Sequence 145, App
34	853.6	29.6	3884	13	US-10-146-728-145	Sequence 145, App
35	853.6	29.6	3884	13	US-10-146-786-145	Sequence 145, App
36	853.6	29.6	3884	13	US-10-147-499-145	Sequence 145, App
37	853.6	29.6	3884	13	US-10-157-798-145	Sequence 145, App
38	853.6	29.6	3884	15	US-10-028-072-145	Sequence 145, App
39	853.6	29.6	3884	15	US-10-121-049-145	Sequence 145, App
40	853.6	29.6	3884	15	US-10-123-904-145	Sequence 145, App
41	853.6	29.6	3884	15	US-10-140-470-145	Sequence 145, App
42	853.6	29.6	3884	15	US-10-175-746-145	Sequence 145, App
43	853.6	29.6	3884	15	US-10-176-918-145	Sequence 145, App
44	853.6	29.6	3884	15	US-10-176-921-145	Sequence 145, App
45	853.6	29.6	3884	15	US-10-137-865-145	Sequence 145, App

#### ALIGNMENTS

#### RESULT 1

US-09-970-944-1  
; Sequence 1, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herrman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shinkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an  
; TITLE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/237,862  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 2881  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-970-944-1

Query Match	100.0%	Score 2881;	DB 11;	Length 2881;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 2881;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	AGCTGGGGCTCCGGGCTGAGCGCTAAAGCGCCCTCCCGCGCGGGCGCCCGCGCGCGG 60		
Db	1	AGCTGGGGCTCCGGGCTGAGCGGTAAAGCGCCCTCCCGCGCGGGCGCCCGCGCGCGG 60		
Qy	61	GCCCGCGCGCTGCGCGCCCGCGCGCCATGCGCGCTCGGCCCGGCGCTGTGGCAGCGCTCC 120		
Db	61	GCCCGCGCGCTGCGCGCCCGCGCGCCATGCGCGCTCGGCCCGGCGCTGTGGCAGCGCTCC 120		
Qy	121	TGGGCATAGTCTCGCGCTTGGCTCCCGCGCTGGGTGCCAGAGTGCCACCGTGG 180		

Db 121 TGGGCAATAGTCTCGCGGCTTGGCTCCGGGCTCGGGTCCCGAGCAGAGTGCCACCGGTGG 180  
Qy 181 COAACCCAGTGTCTGGTGCCAAACCGGACCTCTCTCCCACTTCTCTGGTGAGCCCGAGG 240  
Db 181 COAACCCAGTGTCTGGTGCCAAACCGGACCTCTCTCCCACTTCTCTGGTGAGCCCGAGG 240  
Qy 241 ATGTGTACATCTCTAAGAACAGCCAGTGTCTCTGTGTGCAAGCCGCTGCGCCGACGC 300  
Db 241 ATGTGTACATCTCTAAGAACAGCCAGTGTCTCTGTGTGCAAGCCGCTGCGCCGACGC 300  
Qy 301 AGATCTTCTTCAAGTGTCAACGGGAGTGGGTGGCGCCAGGTGCACACAGTGCATCGACGCA 360  
Db 301 AGATCTTCTTCAAGTGTCAACGGGAGTGGGTGGCGCCAGGTGCACACAGTGCATCGACGCA 360  
Qy 361 GCACAGACGGGAGCGTGTGAGCCGACCATGAGAGTGCAGATTAATGTCTCAAGSCAGC 420  
Db 361 GCACAGACGGGAGCGTGTGAGCCGACCATGAGAGTGCAGATTAATGTCTCAAGSCAGC 420  
Qy 421 AGGTGAGAGAGTGTTCGGGCTGGAGGAATACTGTGCGCAGTGCCTGGCATGAGCTCCT 480  
Db 421 AGGTGAGAGAGTGTTCGGGCTGGAGGAATACTGTGCGCAGTGCCTGGCATGAGCTCCT 480  
Qy 481 CGGGACCAACCAAGAGTCAGAGGCCCTACATCCGATAGCCAGATTGCGCAAGAACTTCG 540  
Db 481 CGGGACCAACCAAGAGTCAGAGGCCCTACATCCGATAGCCAGATTGCGCAAGAACTTCG 540  
Qy 541 AGCAGAGCGGTGTGCGCAAGGAGGTCTCTGAGCAGGGCATCTGTGTCTGCCCTGCCGTC 600  
Db 541 AGCAGAGCGGTGTGCGCAAGGAGGTCTCTGAGCAGGGCATCTGTGTCTGCCCTGCCGTC 600  
Qy 601 CACCGAGAGCATCCCTCCAGCCGAGGTGGGTCCGGAAACAGAGGACCTGGTGGAC 660  
Db 601 CACCGAGAGCATCCCTCCAGCCGAGGTGGGTCCGGAAACAGAGGACCTGGTGGAC 660  
Qy 661 CGTCCCTGGACCCCAATATATACATCACGCGGAGCACAGCCTGGTGGCAGAGCCC 720  
Db 661 CGTCCCTGGACCCCAATATATACATCACGCGGAGCACAGCCTGGTGGCAGAGCCC 720  
Qy 721 GCCTTGCTGACAGGCGCAATACATCTGGTGCGCAGAAACATCTGTGGCAGCTCGCGCA 780  
Db 721 GCCTTGCTGACAGGCGCAATACATCTGGTGCGCAGAAACATCTGTGGCAGCTCGCGCA 780  
Qy 781 GCGCTCCGCTGTGTCATCTGTCTGTAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840  
Db 781 GCGCTCCGCTGTGTCATCTGTCTGTAACCGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 840  
Qy 841 CCGTCTGCAAGCGCAGTGTGGCGCGCTGGCAGAAACGAGCCGAGCTGCACCAACC 900  
Db 841 CCGTCTGCAAGCGCAGTGTGGCGCGCTGGCAGAAACGAGCCGAGCTGCACCAACC 900  
Qy 901 CGGCGCTCTCAACGGGGGCGCTTCTGTGAGGGGCGAGTGTGGTGGTGGTGGTGGTGGT 960  
Db 901 CGGCGCTCTCTCAACGGGGGCGCTTCTGTGAGGGGCGAGTGTGGTGGTGGTGGTGGTGGT 960  
Qy 961 CCTCTCTGTGTCCTGTGAGCGGAGCTGGAGCCGCTGGAGCAAGTGGTGGTGGTGGTGGT 1020  
Db 961 CCTCTCTGTGTCCTGTGAGCGGAGCTGGAGCCGCTGGAGCAAGTGGTGGTGGTGGTGGT 1020  
Qy 1021 GGCTGAGTGCACCCACTGCGGAGCGCTGAGTGTCTCTGACCCAGCACCCCGCAACGGAG 1080  
Db 1021 GGCTGAGTGCACCCACTGCGGAGCGCTGAGTGTCTCTGACCCAGCACCCCGCAACGGAG 1080  
Qy 1081 GGGAGAGTGCAGGGACCTGACCTGACACCCGCACTGTACCAAGTGAAGTCTGTGTAC 1140  
Db 1081 GGGAGAGTGCAGGGACCTGACCTGACACCCGCACTGTGTACCAAGTGAAGTCTGTGTAC 1140  
Qy 1141 ACAGTGTCTTGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATCGCCGTGGCCGTCT 1200  
Db 1141 ACAGTGTCTTGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATCGCCGTGGCCGTCT 1200  
Qy 1201 GCCTGCTCTCTGCTGTCT 1260  
Db 1201 GCCTGCTCTCTGCTGTCT 1260

Qy 1261 ACTCAGATGTGGCTGACTGCTCCATTTCTCACTCAGGTTTCCAGCCCGTTCAGCATCAAGC 1320  
Db 1261 ACTCAGATGTGGCTGACTGCTCCATTTCTCACTCAGGTTTCCAGCCCGTTCAGCATCAAGC 1320  
Qy 1321 CCAGCAAGAGCAGCAACCCCATCTCTGCTCACCATCCAGCCGACCTTCAGCACACCA 1380  
Db 1321 CCAGCAAGAGCAGCAACCCCATCTCTGCTCACCATCCAGCCGACCTTCAGCACACCA 1380  
Qy 1381 CCTTACAGGGCAGTCTGTGTCCCGCAGGATGGGCCAGCCCGAGTTCCAGCTCACCA 1440  
Db 1381 CCTTACAGGGCAGTCTGTGTCCCGCAGGATGGGCCAGCCCGAGTTCCAGCTCACCA 1440  
Qy 1441 ATGGGCACTGTCTCAGCCCTTGGGTGGCGCGCCACACACTGTCACCAACAGCTCTCCCA 1500  
Db 1441 ATGGGCACTGTCTCAGCCCTTGGGTGGCGCGCCACACACTGTCACCAACAGCTCTCCCA 1500  
Qy 1501 CCTCTGAGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560  
Db 1501 CCTCTGAGCCGAGGAGTTCGTCTCCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGC 1560  
Qy 1561 CCGGAGGACACAGCAACATGACTATGGACCTTCAACTTCTCGGGGCGGCTCATGA 1620  
Db 1561 CCGGAGGACACAGCAACATGACTATGGACCTTCAACTTCTCGGGGCGGCTCATGA 1620  
Qy 1621 TCCCTAATACAGTATCAGCCTCTCATCCCGCAGATGCCATACCCCGAGGAAAGATCT 1680  
Db 1621 TCCCTAATACAGTATCAGCCTCTCATCCCGCAGATGCCATACCCCGAGGAAAGATCT 1680  
Qy 1681 ATGAGATCTACCTCAGCTGCAAGCCGAAAGACGTGAGTTGGCTTGGCTTGGCTGTC 1740  
Db 1681 ATGAGATCTACCTCAGCTGCAAGCCGAAAGACGTGAGTTGGCTTGGCTTGGCTGTC 1740  
Qy 1741 AGACCTCTGAGTCCCATCTGTTAGCTGTGACCCCTCGGCTCTCAGCCGCGCAG 1800  
Db 1741 AGACCTCTGAGTCCCATCTGTTAGCTGTGACCCCTCGGCTCTCAGCCGCGCAG 1800  
Qy 1801 TCATCTCTGCTATGAGACCACTGTGGGAGCCAGCCCTGACAGCTGAGGCTTCGCTCA 1860  
Db 1801 TCATCTCTGCTATGAGACCACTGTGGGAGCCAGCCCTGACAGCTGAGGCTTCGCTCA 1860  
Qy 1861 AAAAGCAGTCTGTCGAGGGCAGCTGGGAGCAGGATGTGTGACCTTGGGCGAGAGGCGC 1920  
Db 1861 AAAAGCAGTCTGTCGAGGGCAGCTGGGAGCAGGATGTGTGACCTTGGGCGAGAGGCGC 1920  
Qy 1921 CTTCCCACTCTACTACTGCGCAGCTGGAGCCAGTCCCTGCTACGCTTTCACCGAGCAGC 1980  
Db 1921 CTTCCCACTCTACTACTGCGCAGCTGGAGCCAGTCCCTGCTACGCTTTCACCGAGCAGC 1980  
Qy 1981 TGGGCGCTTTGGCTTGGTGGGAGAGCCCTCAGCGTGGCTGGCGCAAGCGCTCAAGC 2040  
Db 1981 TGGGCGCTTTGGCTTGGTGGGAGAGCCCTCAGCGTGGCTGGCGCAAGCGCTCAAGC 2040  
Qy 2041 TGCTTCTGTTTGGCGCGGTGCGACCTTCCCTCGAGTACAACTCCGGGTCTACTGCC 2100  
Db 2041 TGCTTCTGTTTGGCGCGGTGCGACCTTCCCTCGAGTACAACTCCGGGTCTACTGCC 2100  
Qy 2101 TGCAATGACACCCAGATGCACTCAAGAGGTGTGAGCTGGAGAGACAGCTGGGGGAC 2160  
Db 2101 TGCAATGACACCCAGATGCACTCAAGAGGTGTGAGCTGGAGAGACAGCTGGGGGAC 2160  
Qy 2161 AGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTACCAACACTTCGCTAT 2220  
Db 2161 AGCTGATCCAGGAGCCACGGGTCTGCACTTCAAGGACAGTTACCAACACTTCGCTAT 2220  
Qy 2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGATGAGCTCTTGTGAGTACACAGAGA 2280  
Db 2221 CCATCCAGATGTGCCAGCTCCCTGTGGAAGATGAGCTCTTGTGAGTACACAGAGA 2280  
Qy 2281 TCCCTTTTATACATCTGGAATGGCACGCGGTACTTGTGACTGACCTTCAACCTGG 2340  
Db 2281 TCCCTTTTATACATCTGGAATGGCACGCGGTACTTGTGACTGACCTTCAACCTGG 2340



QY 2341 AGCGTGTAGCCCAACACCTAGTACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGAGG 2400  
Db |||||  
QY 2341 AGCGTGTAGCCCAACACCTAGTACCTGGCTGCAAGCTGTGGGTGTGGCAGGTGAGG 2400  
Db |||||  
QY 2401 GCGACGGGCGAGCTTCAGCATCACTTCAACATCAACCAAGGACACAAGGTTTGTGAGC 2460  
Db |||||  
QY 2401 GCGACGGGCGAGCTTCAGCATCACTTCAACATCAACCAAGGACACAAGGTTTGTGAGC 2460  
Db |||||  
QY 2461 TGTGCTGCTTGGAGAGTGAAGCGGGGTGCCAGCCTGTGGTGGGCCCAAGTTCCTTCAAG 2520  
Db |||||  
QY 2461 TGTGCTGCTTGGAGAGTGAAGCGGGGTGCCAGCCTGTGGTGGGCCCAAGTTCCTTCAAG 2520  
Db |||||  
QY 2521 TCCCTTCTCTCATTCCGCGAAGATTAATTTCCAGCCTGGACCCACCTGTAGCGGGGTG 2580  
Db |||||  
QY 2521 TCCCTTCTCTCATTCCGCGAAGATTAATTTCCAGCCTGGACCCACCTGTAGCGGGGTG 2580  
Db |||||  
QY 2581 CCGACTGGCGGACTTGGCCCGAAGAACTCCACTGGACAGCCATCTCAGCTTCTTTGCC 2640  
Db |||||  
QY 2581 CCGACTGGCGGACTTGGCCCGAAGAACTCCACTGGACAGCCATCTCAGCTTCTTTGCC 2640  
Db |||||  
QY 2641 CCAAGCCCAAGCCCAAGCAGCATGATCTCAACCTGTGGAGCGCGGCACCTTCCCAAG 2700  
Db |||||  
QY 2641 CCAAGCCCAAGCCCAAGCAGCATGATCTCAACCTGTGGAGCGCGGCACCTTCCCAAG 2700  
Db |||||  
QY 2701 GCAACTCAGCAGTGGCTGAGCAGTGGCTGAGCTGGAGCTGGCTGGAGCTGGCTGGAG 2760  
Db |||||  
QY 2701 GCAACTCAGCAGTGGCTGAGCAGTGGCTGAGCTGGAGCTGGCTGGAGCTGGCTGGAG 2760  
Db |||||  
QY 2761 TCACAGTGTGGAGCTGAGTGTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2820  
Db |||||  
QY 2761 TCACAGTGTGGAGCTGAGTGTGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2820  
Db |||||  
QY 2821 CTTTGGCCCAACCAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880  
Db |||||  
QY 2821 CTTTGGCCCAACCAAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2880  
Db |||||  
QY 2881 A 2881  
Db |||||  
QY 2881 A 2881  
Db |||||

## RESULT 2

US-09-918-779-1  
; Sequence 1, Application US/0918779  
; Publication No. US20030064369A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier, Raymond  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Grosse, William  
; APPLICANT: Alsobrook, John  
; APPLICANT: Lepley, Denise  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David  
; APPLICANT: Smithson, Glennda  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-074 US  
; CURRENT APPLICATION NUMBER: US/09/918,779  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/221,409  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/222,840  
; PRIOR FILING DATE: 2000-08-04

QY 42 CGCGGGGCG 101  
Db 1 CGCGGGGCG 60  
QY 102 GGCCTGTGCGCAGCGCTCTTGGGCATAGTCTCGCGCGCTTGGCTCGCGCGCTCGCGGTGC 161  
Db 61 GGCCTGTGCGCAGCGCTCTTGGGCATAGTCTCGCGCGCTTGGCTCGCGCGCTCGCGGTGC 120  
QY 162 CAGCAGAGTGCACCGTGGCCAAACCCAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 221  
Db 121 CAGCAGAGTGCACCGTGGCCAAACCCAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTG 180  
QY 222 TTCCTGTGAGCGCGCGAGGATGTGTATCTGTCAAGAAACAAGCCAGTGTGCTGTGCTG 281  
Db 181 TTCCTGTGAGCGCGCGAGGATGTGTATCTGTCAAGAAACAAGCCAGTGTGCTGTGCTG 240  
QY 282 AAGCGCGTGGCG 341  
Db 241 AAGCGCGTGGCG 300  
QY 342 GACCACGTGTATCGAGCGCGCAGCAGACGCGGAGCAGTGTGCTGAGCGCGCGCGCGCG 401  
Db 301 GACCACGTGTATCGAGCGCGCAGCAGACGCGGAGCAGTGTGCTGAGCGCGCGCGCGCG 360  
QY 402 ATTAATGTCTCAAGCGCAGCGTGTGAGAAAGTGTTCGGCGCTGGAGGATATCTGTGCG 461  
Db 361 ATTAATGTCTCAAGCGCAGCGTGTGAGAAAGTGTTCGGCGCTGGAGGATATCTGTGCG 420  
QY 462 TGCCTGCGCATGAGCTCTCGGGCACCACCAAGAGTGTGAGAGGCTATCATCCGATAGCC 521  
Db 421 TGCCTGCGCATGAGCTCTCGGGCACCACCAAGAGTGTGAGAGGCTATCATCCGATAGCC 480  
QY 522 AGATTGCGCAAGAACTTTCGAGCAGGAGCGCTGCGCCCAAGAGGAGTGTCCCTGGAGCAG 581  
Db 481 AGATTGCGCAAGAACTTTCGAGCAGGAGCGCTGCGCCCAAGAGGAGTGTCCCTGGAGCAG 540  
QY 582 ATCGTGTGCGCTGCGCGTCCACCGAGGAGGATCCCTCCAGCGCGAGGTGGAGTGGCTCGG 641  
Db 541 ATCGTGTGCGCTGCGCGTCCACCGAGGAGGATCCCTCCAGCGCGAGGTGGAGTGGCTCGG 600  
QY 642 AACGAGACCTGTGTGGACCCCGTCCCTGGACCCCAATGTATATCATCATCGCGGAGCAGC 701

Query Match 92.9%; Score 2676.4; DB 13; Length 2752;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 2728; Conservative 0; Mismatches 21; Indels 9; Gaps 3;

Db	601	AA	CGAGGAC	CTG	TGGAC	CCG	TCC	TGG	ACCA	TG	TAT	CAT	CAC	GCG	GAG	CAC	G	660
Qy	702	CT	GCTGGT	TGGA	CAG	GGC	CGC	CTT	GT	GTG	A	CAG	GC	CA	ACT	A	G	761
Db	661	CT	GCTGGT	TGGA	CAG	GGC	CGC	CTT	GT	GTG	A	CAG	GC	CA	ACT	A	G	720
Qy	762	AT	GCTGG	CAC	GTG	CGC	GAC	GCG	CTC	CG	TGT	GT	CT	AT	C	GT	T	821
Db	721	AT	GCTGG	CAC	GTG	CGC	GAC	GCG	CTC	CG	TGT	GT	CT	AT	C	GT	T	780
Qy	822	TC	GAC	TGG	AC	CG	AG	TGG	T	CCG	CT	G	CAG	CG	CC	AG	CT	881
Db	781	TC	GAC	TGG	AC	CG	AG	TGG	T	CCG	CT	G	CAG	CG	CC	AG	CT	840
Qy	882	AG	CGG	AG	CTG	C	AC	AA	CC	CGG	CG	CTT	C	T	CA	A	CG	941
Db	841	AG	CGG	AG	CTG	C	AC	AA	CC	CGG	CG	CTT	C	T	CA	A	CG	900
Qy	942	GT	CC	AT	CAC	CG	C	CG	T	CC	T	CT	G	T	G	T	G	1001
Db	901	GT	CC	AG	AA	---	AA	C	AG	CT	G	GC	A	CC	T	G	T	957
Qy	1002	AG	CA	A	G	TGG	T	G	G	C	T	G	G	C	T	G	G	1061
Db	958	AG	CA	A	G	TGG	T	G	G	C	T	G	G	C	T	G	G	1017
Qy	1062	CC	AG	C	CC	CG	CA	AC	CG	AG	G	G	AG	G	AG	T	G	1121
Db	1018	CC	AG	C	CC	CG	CA	AC	CG	AG	G	G	AG	G	AG	T	G	1077
Qy	1122	AC	CA	G	T	G	A	C	T	G	T	G	A	C	T	G	A	1181
Db	1078	AC	CA	G	T	G	A	C	T	G	T	G	A	C	T	G	A	1137
Qy	1182	CT	CA	T	CG	CG	TGG	CG	CT	G	T	G	T	CT	CT	CA	T	1241
Db	1138	CT	CA	T	CG	CG	TGG	CG	CT	G	T	G	T	CT	CT	CA	T	1197
Qy	1242	CG	GA	AG	AG	G	G	G	G	T	G	G	A	C	T	G	A	1301
Db	1198	CG	GA	AG	AG	G	G	G	G	T	G	G	A	C	T	G	A	1257
Qy	1302	CAG	CC	CG	T	CAG	CAT	CA	A	G	CC	CA	AG	C	A	A	C	1361
Db	1358	CAG	CC	CG	T	CAG	CAT	CA	A	G	CC	CA	AG	C	A	A	C	1317
Qy	1362	GAC	CT	CAG	---	C	A	C	C	A	C	C	A	C	C	A	C	1418
Db	1318	GAC	CT	CAG	C	C	A	C	C	A	C	C	A	C	C	A	C	1377
Qy	1419	AG	C	C	C	C	A	G	T	T	C	A	G	G	C	A	G	1478
Db	1378	AG	C	C	C	C	A	G	T	T	C	A	G	G	C	A	G	1437
Qy	1479	AC	AT	G	C	A	C	A	G	T	T	C	A	G	G	C	A	1538
Db	1438	AC	AT	G	C	A	C	A	G	T	T	C	A	G	G	C	A	1497
Qy	1539	CAG	AA	C	T	A	C	T	T	CG	C	T	CG	C	C	C	G	1598
Db	1498	CAG	AA	C	T	A	C	T	T	CG	C	T	CG	C	C	C	G	1557
Qy	1599	TT	CT	CG	G	G	G	G	G	C	T	G	A	T	C	CA	T	1658
Db	1558	TT	CT	CG	G	G	G	G	G	C	T	G	A	T	C	CA	T	1617
Qy	1659	GC	CA	T	AC	CC	CG	AG	GA	AG	AT	CT	AT	G	A	G	A	1718
Db	1618	GC	CA	T	AC	CC	CG	AG	GA	AG	AT	CT	AT	G	A	G	A	1677
Qy	1719	AG	GT	T	CG	CC	CT	AG	T	GG	T	G	T	G	T	G	A	1778



```
QY 1362 GACCTCAG---CACACACACACCTACAGGGCAGTCTCTGTCCCGCGCAGGATGGGCC 1418
DB 1318 GACCTCAGCACCACACACACCTACAGGGCAGTCTCTGTCCCGCGCAGGATGGGCC 1377
QY 1419 AGCCCAAGTTCCAGCTCACCAATGGGACCTGTCTCAGCCCTCGGTGGCGCGCCAC 1478
DB 1378 AGCCCAAGTTCCAGCTCACCAATGGGACCTGTCTCAGCCCTCGGTGGCGCGCCAC 1437
QY 1479 ACACCTGCACACAGCTCTCCACCTCTGAGGCCGAGGATTCGTCTCCGCCCTCTCCACC 1538
DB 1438 ACACCTGCACACAGCTCTCCACCTCTGAGGCCGAGGATTCGTCTCCGCCCTCTCCACC 1497
QY 1539 CAGAACTACTTCGGCTCCCTGCCCGAGCAGCAGCAATGATGGGACCTTCAAC 1598
DB 1498 CAGAACTACTTCGGCTCCCTGCCCGAGCAGCAGCAATGATGGGACCTTCAAC 1557
QY 1599 TTCTCTGGGGCGCGCTGATGATCTCTAATACAGGTATCAGCTCTCTCATCCCGCCAGAT 1658
DB 1558 TTCTCTGGGGCGCGCTGATGATCTCTAATACAGGTATCAGCTCTCTCATCCCGCCAGAT 1617
QY 1659 GCCATACCCCGAGGAGGATCTATGAGATCTACCTCAGCTGCACAAAGCGGAAGACGTG 1718
DB 1618 GCCATACCCCGAGGAGGATCTATGAGATCTACCTCAGCTGCACAAAGCGGAAGACGTG 1677
QY 1719 AGTTGGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCGTTAGCTGTGACCCCT 1778
DB 1678 AGTTGGCCCTAGCTGGCTGTGAGACCTGTGAGTCCCATCGTTAGCTGTGACCCCT 1737
QY 1779 GCGTCTCTCTCACCCGCGCAGTCATCTCTGGGTATGAGCACTGTGGGAGCCGACCCCT 1838
DB 1738 GCGTCTCTCTCACCCGCGCAGTCATCTCTGGGTATGAGCACTGTGGGAGCCGACCCCT 1797
QY 1839 GACAGCTGAGCTGCGCTCAAAAAGCAGTCTGCGAGGGCAGCTGGAGCAGGATGTG 1898
DB 1798 GACAGCTGAGCTGCGCTCAAAAAGCAGTCTGCGAGGGCAGCTGGG---AGGATGTG 1854
QY 1899 GTGACCTGGGAGAGGGGCCCTCCACCTTACTACTGCTGAGCTGAGGAGGAGCCATGSC 1958
DB 1855 GTGACCTGGGAGAGGGGCCCTCCACCTTACTACTGCTGAGCTGAGGAGGAGCCATGSC 1914
QY 1959 TGTAGCTTCTACCCAGCAGCTGGGCGCTTTGCGCTGTGGAGAGGCGCTTCACGCTG 2018
DB 1915 TGTAGCTTCTACCCAGCAGCTGGGCGCTTTGCGCTGTGGAGAGGCGCTTCACGCTG 1974
QY 2019 GCTGCCGCCAAGCGCTCAAGCTGCTTTGTCGCGCGGTGGCTGCACCTCCCTCGAG 2078
DB 1975 GCTGCCGCCAAGCGCTCAAGCTGCTTTGTCGCGCGGTGGCTGCACCTCCCTCGAG 2034
QY 2079 TACAACTCGGCTTACTGCTGATGACACCCAGATGCTCAAGGAGGTGGTGCAG 2138
DB 2035 TACAACTCGGCTTACTGCTGATGACACCCAGATGCTCAAGGAGGTGGTGCAG 2094
QY 2139 CTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCCAGCGGTCTCTGCACTTCAAGGAC 2198
DB 2095 CTGGAGAAGCAGCTGGGGGACAGCTGATCCAGAGCCAGCGGTCTCTGCACTTCAAGGAC 2154
QY 2199 AGTTACCAACCTGGCCCTTATCCATCCAGATGTGCCAGCTGCTCTGTGGAGAGATAG 2258
DB 2155 AGTTACCAACCTGGCCCTTATCCATCCAGATGTGCCAGCTGCTCTGTGGAGAGATAG 2214
QY 2259 CTCTTGTGAGCTACAGGAGATCCCTTTTATCATCTGGAATGCGACGCGGTAC 2318
DB 2215 CTCTTGTGAGCTACAGGAGATCCCTTTTATCATCTGGAATGCGACGCGGTAC 2274
QY 2319 TTGCACTGCACTTACCCCTGGAGCGTGTGAGCCCGCAGCACTAGTACCTGGCTGCAAG 2378
DB 2275 TTGCACTGCACTTACCCCTGGAGCGTGTGAGCCCGCAGCACTAGTACCTGGCTGCAAG 2334
QY 2379 CTGTGGGTGTGGAGGTGGAGGGCGAGGGCAGAGCTTACAGTCACTTCAACATCACC 2438
DB 2335 CTGTGGGTGTGGAGGTGGAGGGCGAGGGCAGAGCTTACAGTCACTTCAACATCACC 2394
```

```
QY 2439 AAGGACACAAGTTTGTGAGCTGTGCTCTGGAGAGTGAAGCGGGGTCCACGCCCTG 2498
DB 2395 AAGGACACAAGTTTGTGAGCTGTGCTCTGGAGAGTGAAGCGGGGTCCACGCCCTG 2454
QY 2499 GTGGGCCCAAGTTCCTTCAAGATCCCTTCTCATTTGGCAGAGATATTTCCAGCCTG 2558
DB 2455 GTGGGCCCAAGTTCCTTCAAGATCCCTTCTCATTTGGCAGAGATATTTCCAGCCTG 2514
QY 2559 GACCCACACCTGTAGCGGGGTGCGACTGGGGGACTCTGGGCCAGAAACTCCACCTGGAC 2618
DB 2515 GACCCACACCTGTAGCGGGGTGCGACTGGGGGACTCTGGGCCAGAAACTCCACCTGGAC 2574
QY 2619 AGCCATCTCAGCTTCTTTGCTCTCAAGCCCGAGCCCGCAGCCATGATCTCAACTGTGG 2678
DB 2575 AGCCATCTCAGCTTCTTTGCTCTCAAGCCCGAGCCCGCAGCCATGATCTCAACTGTGG 2634
QY 2679 GAGGCGGGCAGCTTCCCAACCGGAACTCAGCAGCTGGCTGCAGCAGTGGCTGGACTG 2738
DB 2635 GAGGCGGGCAGCTTCCCAACCGGAACTCAGCAGCTGGCTGCAGCAGTGGCTGGACTG 2694
QY 2739 GGCCAGCCAGACGCTGGCTCTTTCAGAGTGTGCGAGGCTGAGTGTGAGGCGCGCCAG 2796
DB 2695 GGCCAGCCAGACGCTGGCTCTTTCAGAGTGTGCGAGGCTGAGTGTGAGGCGCGCCAG 2752
```

## RESULT 4

```
US-10-311-623-13
; Sequence 13, Application US/10311623
; Publication No. US20040023244A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.
; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.
; APPLICANT: YUE, Henry; NGUYEN, Dannel B.
; APPLICANT: TANG, Y. Tom; LAL, Preeti G.
; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Yalda
; APPLICANT: LU, Dying Aina M.; GRAUL, Richard C.
; APPLICANT: YAO, Monique G.; BURFORD, Neil
; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.
; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.
; APPLICANT: YANG, Junming; XU, Yuming
; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.
; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.
; APPLICANT: DUGGAN, Brendan M.; LU, Yan
; TITLE OF INVENTION: RECEPTORS
; FILE REFERENCE: PF-0793 USN
; CURRENT APPLICATION NUMBER: US/10/311,623
; PRIOR FILING DATE: 2002-12-17
; PRIOR APPLICATION NUMBER: US 01/19942
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/214,027
; PRIOR FILING DATE: 2000-06-21
; PRIOR APPLICATION NUMBER: US 60/228,045
; PRIOR FILING DATE: 2000-08-25
; PRIOR APPLICATION NUMBER: US 60/255,104
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 3580
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CB1
US-10-311-623-13
```

Query Match 83.4%; Score 2402.6; DB 17; Length 3580;  
Best Local Similarity 93.4%; Pred. No. 0;  
Matches 2615; Conservative 0; Mismatches 9; Indels 177; Gaps 3;

```
QY 84 GCCATGCGCGCTCGCGCCCGCGCCCTGTGCGCAGCGCTCTCTGCGCATAGTCTTCGCGCTGG 143
DB 1 GCCATGCGCGCTCGCGCCCGCGCCCTGTGCGCAGCGCTCTCTGCGCATAGTCTTCGCGCTGG 60
```



2047 AATGACGCGAGCGTACTTGACATGACCTTACCTTGAGCGTGTAGCCCGAGCACT 2106  
2361 AGTGACCTGGCGCTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGACGGGCGAGGCTTCAGC 2420  
2107 AGTGACCTGGCGCTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGACGGGCGAGGCTTCAGC 2166  
2421 ATCACTTCAACATCACCAAGACACACAGGTTTGTGAGCTGTGAGTCTGTGAGAGTGAA 2480  
2167 ATCACTTCAACATCACCAAGACACACAGGTTTGTGAGCTGTGAGTCTGTGAGAGTGAA 2226  
2481 GCGGGGTCCAGCGCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAG 2540  
2227 GCGGGGTCCAGCGCTGTGGGCGCCAGTGTCTTCAAGATCCCTTCTCATTCGGCAG 2286  
2541 AAGATAATTTCCAGCGCTGACCCACCTGTAGCGGGGTGCGGACTTGGCGGACTTGGCC 2600  
2287 AAGATAATTTCCAGCGCTGACCCACCTGTAGCGGGGTGCGGACTTGGCGGACTTGGCC 2346  
2601 CAGAACTCCAGCTGAGAGCCATCTCAGCTTCTTGTGCTTCCAGCCAGCCCGCAGCC 2660  
2347 CAGAACTCCAGCTGAGAGCCATCTCAGCTTCTTGTGCTTCCAGCCAGCCCGCAGCC 2406  
2661 ATGATCTCAACTGTGGAGGGCGGCACTTCCCAACCGCAACCTCAGCCAGCTGGCT 2720  
2407 ATGATCTCAACTGTGGAGGGCGGCACTTCCCAACCGCAACCTCAGCCAGCTGGCT 2466  
2721 GCAGCAGTGGCTGAGCTGGGCGCAGCAGCGTGGCTCTTACAGTGTGGAGGCTGAG 2780  
2467 GCAGCAGTGGCTGAGCTGGGCGCAGCAGCGTGGCTCTTACAGTGTGGAGGCTGAG 2526  
2781 TGTGAGGCGGCGCAGGCGCGACCTACACTCTCAGCAGCTTGGCAGCCACCAAGGAC 2840  
2527 TGTGAGGCGGCGCAGGCGCGACCTACACTCTCAGCAGCTTGGCAGCCACCAAGGAC 2586  
2841 AGGCAAGCGCGACAGGGGCGCTTCCCAACCGGGGAGA 2881  
2587 AGGCAAGCGCGACAGGGGCGCTTCCCAACCGGGGAGA 2627

RESULT 5  
US-09-933-261-1  
Sequence 1, Application US/09933261  
Publication No. US20030040046A1  
GENERAL INFORMATION:  
APPLICANT: Tessier-Lavigne, Marc  
Leonardo, E. David  
Hink, Lindsey  
Masu, Masayuki  
Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/933,261  
FILING DATE: 20-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/808,982  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A

REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3014 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-933-261-1  
Query Match 77.6%; Score 2235.8; DB 10; Length 3014;  
Best Local Similarity 88.5%; Pred. No. 0;  
Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;  
QY 87 ATGGCGGCTCGGGCGCGGCTGTGGCCAGCGCTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 146  
Db 1 ATGGCGGCTCGGGCGCGGCTGTGGCCAGTGTCTCTGGGCATAGTCTCTCGCGCTTGGCTC 60  
QY 147 CCGCGCTCGGGTCCCGCAGCAGAGTGCACCGTGGCCCAACCCAGTGTCTGTGTCACACCG 206  
Db 61 CGTGGTTCGGGTGCCCGCAGCAGAGTGCACCGTGGCCCAATCCAGTGTCCCGTGTCCACCC 120  
QY 207 GACCTGCTTCCCACTTCTCTGTGGAGCCCGGAGGATGTGTAGTCTCTCAAGAACAAAGCCA 266  
Db 121 GACCTGCTTCCCACTTCTCTGTGGAGCCCGGAGGATGTGTAGTCTCTCAAGAACAAAGCCA 180  
QY 267 GTGCTGCTTGTGTGCAAGGCGCTGCGCCAGCAGATCTTCTTCAAGTGCACACGGGAG 326  
Db 181 GTGTTGTTGTGTGCAAGGCTGTGCTGCTGCCACCCAGATCTTCTTCAAGTGCATGGGAA 240  
QY 327 TGGGTGCGCAAGTGTGACCAACCGTGTGAGCGCGCAGCAGACGGGAGCAGTGTGTAGCCG 386  
Db 241 TGGGTGCGCAGGTGCATCAGCTAATTGAACGCGCAGCAGCCAGCAGCAGCGGATTTGCCA 300  
QY 387 ACCATGGAGTCCGCAATTAATGTCTCAAGCGCAGCGTGTGAGAGGTGTTCGGGCTGGAG 446  
Db 301 ACCATGGAGTCCGATATCAACGTATCGAGCGCAGGTGTAGAGAGTGTTCGGGCTGGAG 360  
QY 447 GAATCTGCTGCGAGTGTGCGTGTGCGATGAGCTCTCTCGGGCACCACCAAGAGTCAAGAGGCC 506  
Db 361 GAATCTGCTGCGAGTGTGCGATGAGCTCTCTCGGGTACCAACAAAGTCAAGAGGCC 420  
QY 507 TACATCCGATAGCAGATTCGCAAGAACTTCGAGCAGGAGCGCTGTGCGCAAGGAGGTG 566  
Db 421 TACATCCGATAGCTTATTTGCGCAAGAACTTTGAGCAGGAGCCACTGGGCCAAGGAGTG 480  
QY 567 TCCCTGGAGCGGCGATCGTGTGCTGCGCTCCACCGGAGGCGATCCCTCCAGCGCGAG 626  
Db 481 TCATCGAGCAGGCGATTTACTTACTTGTGCGCGCCCGCAGAGAGATTCCTCCAGCTGAG 540  
QY 627 GTGAGTGTGCTCCGGAACGAGGACCTGTGGTGGACCCCGTCCCTGGACCCCAATATATCATC 686  
Db 541 GTGAGTGTGCTTCAATGAGGACCTGTGGACCCCTCGTGGACCCCTCCCTCCATGTTATCATC 600  
QY 687 ACSCGGGAGCAGCGCTGTGGTGTGAGCAGGCGCGGCTTCTGACACGGCCAACTACACC 746  
Db 601 ACSCGGGAGCAGCGCTGTGGTGTGAGCAGGCGCGGCTTGGCCGACACGGCCAACTACACC 660  
QY 747 TCGGTGCGCAAGAACATCGTGGCAGCTGTGCGCGCAGCGCGCTCCGCTGTGTCTATCTGTAC 806  
Db 661 TGTGTGCGCAAGAACATCGTAGCCGTCGCCGCGAGGACCTCTGCGAGGGTCAATGTTTAT 720  
QY 807 GTGAACGGTGTGGTGTGAGCTGTGAGCGAGTGTGTCCGTGTGACGCGCAGCTGTGGGCGC 866  
Db 721 GTGAACGGTGTGGTGTGAGCTGTGAGCTGTGTCCGTGTGACGCGCAGCTGTGGGCGT 780  
QY 867 GGCTGGCAGAAACGGAGCGGAGGTGCACCAACCGCGCGCTCTCAACGGGGGCGCTTC 926





;  
 ; NUMBER OF SEQUENCES: 8  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 ; STREET: 268 BUSH STREET, SUITE 3200  
 ; CITY: SAN FRANCISCO  
 ; STATE: CALIFORNIA  
 ; COUNTRY: USA  
 ; ZIP: 94104  
 ;  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ;  
 ; CURRENT APPLICATION DATA: US/10/256,702  
 ; APPLICATION NUMBER: US/10/256,702  
 ; FILING DATE: 27-Sep-2002  
 ; CLASSIFICATION: <Unknown>  
 ;  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/933,261  
 ; FILING DATE: 20-Aug-2001  
 ; APPLICATION NUMBER: 08/808,982  
 ; FILING DATE: <Unknown>  
 ;  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: OSMAN, RICHARD A  
 ; REGISTRATION NUMBER: 36,627  
 ; REFERENCE/DOCKET NUMBER: UC96-217  
 ;  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (415) 343-4341  
 ; TELEFAX: (415) 343-4342  
 ;  
 ; INFORMATION FOR SEQ ID NO: 1:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 3014 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ;  
 ; MOLECULE TYPE: cDNA  
 ;  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 ;  
 ; US-10-256-702-1  
 ;  
 ;  
 ; Query Match 77.6%; Score 2235.8; DB 15; Length 3014;  
 ; Best Local Similarity 88.5%; Pred. No. 0;  
 ; Matches 2474; Conservative 0; Mismatches 312; Indels 11; Gaps 4;  
 ;  
 ;  
 ; 87 ATGGCGCTCGGCGCGGCTGGCCAGAGTGCACCGTGGCCAAACCCAGTGGCTGGTGCACACCGG 206  
 ; 1 ATGGCGCTCGGCGCGGCTGGCCAGAGTGCACCGTGGCCAAACCCAGTGGCTGGTGCACACCGG 60  
 ;  
 ; 147 CGCGCTCGGCTCGGCGCGGCTGGCCAGAGTGCACCGTGGCCAAACCCAGTGGCTGGTGCACACCGG 206  
 ; 61 CGTGGTTCGGTGGCGGCGGCTGGCCAGAGTGCACCGTGGCCAAACCCAGTGGCTGGTGCACACCGG 120  
 ;  
 ; 207 GACCTGCTCCGCACTTCCTGGTGGAGCGCGAGATGTGTATCGTCAAGAAACAGGCCA 266  
 ; 121 GACCTGCTCCGCACTTCCTGGTGGAGCGCTGGAGCGTGTATCGTCAAGAAACAGGCCG 180  
 ;  
 ; 267 GTGCTCTGTGTGCAAGCGCGGCGCGCCAGCAGATCTTCTCAAGTGCACACCGGAG 326  
 ; 181 GTGTTGTGTGTGCAAGCGTGTGCTGCGCCAGATCTTCTCAAGTGCACACCGGAG 240  
 ;  
 ; 327 TGGGTGCGCAGGTGGACCACTGATCGAGCGCAGCAGCAGCGGAGCGAGTGGTGGAGCGG 386  
 ; 241 TGGGTGCGCAGGTGGACCACTGATCGAGCGCAGCAGCAGCGGAGCGAGTGGTGGAGCGG 300  
 ;  
 ; 387 ACCATGAGTCCGCACTTAAATGTCTAAGCAGCAGGTTCAGAGGTGTTCGGGCTGGAG 446  
 ; 301 ACCATGAGTCCGCACTTAAATGTCTAAGCAGCAGGTTCAGAGGTGTTCGGGCTGGAG 360  
 ;  
 ; 447 GAATCTGTGTGCAAGCGTGTGCGATGGAGTCTCTCGGCGCAGCAGCAGCGGAGCGG 506  
 ; 361 GAATCTGTGTGCAAGCGTGTGCGATGGAGTCTCTCGGCGCAGCAGCAGCGGAGCGG 420  
 ;  
 ; 507 TACATCGCATAGCAGATTCGCAAGAACTTCGAGCAGGAGCGGCTGGCCAAAGGAGTG 566

Db 421 TACATCGCATAGCAGATTCGCAAGAACTTCGAGCAGGAGCCACTGGCCAAAGGAGTG 480  
 Qy 567 TCCCTGGAGCAGGCGATCGTGTGCCCTGCCCTCCACCGGAGGCGATCCCTCAGCCGAG 626  
 Db 481 TCACCTGGAGCAGGCGATTCGCTGTGCGCCCGCCAGAGGAATCCCCAGGCTGAG 540  
 Qy 627 GTGAGTGGCTCCGGAACGAGGACCTGGTGGACCGCTCCCTGGAGCCCAATATATACATC 686  
 Db 541 GTGAGTGGCTTCGAAATGAGGACCTCGTGGACCCCTCCCTCGATCCCAATGTATATC 600  
 Qy 687 AGCGGAGCAGCAGCTGGTGGAGCAGGCGCGCTCTGTGACAGCGGCGCACTACACC 746  
 Db 601 AGCGGAGCAGCAGCTGGTGGAGCAGGCGCGCTCTGTGACAGCGGCGCACTACACC 660  
 Qy 747 TCGGTGGCCAGAAACATCGTGGACGCTCGCGCAGCGCTCGCTGTCTCATGCTATC 806  
 Db 661 TGTGTGGCCAGAAACATCGTGGACCGCTCGCGCAGCGCTCTCATGCTATGTTTAT 720  
 Qy 807 GTGAACGGTGGTGGTTCGACGTGGACCGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 866  
 Db 721 GTGAACGGTGGTGGTTCGACGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 780  
 Qy 867 GGCTGGCAGAAACGAGCGCGGAGCTGCACCAACCGCGCGCTCTCAACGGGCGGCTTTC 926  
 Db 781 GGCTGGCAGAAACGAGCGCGGAGCTGCACCAACCGCGCGCTCTCAACGGGCGGCTTTC 840  
 Qy 927 TGTGAGGGGCGAAGTGCATGACCGCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 986  
 Db 841 TGTGAGGGGCGAAGTGCATGACCGCAGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 897  
 Qy 987 AGCTGGAGCGCGTGGAGCAAGTGGTGGCGCTGGTGGCTGGAGTGCACCCACTGGCGAGC 1046  
 Db 898 AGCTGGAGTTCGTGGAGTAAGTGGTGGCGCTGGTGGCTGGAGTGCACCCACTGGCGAGC 957  
 Qy 1047 CGTGGAGTGTCTGACCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1106  
 Db 958 CGCGAGTGTCTGACCCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1017  
 Qy 1107 GACCCCGCACTGTACAGTGCATCTGTGTACACAGTGTCTTCTGGCGCTTCGAGGACGTG 1166  
 Db 1018 GACCCCGCACTGTACAGTGCATCTGTGTACACAGTGTCTTCTGGCGCTTCGAGGACGTG 1077  
 Qy 1167 GCGCTCTATGTGGCGCTCATCGCGCTGGTGGCGCTGTGCTGGTGGTGGTGGTGGTGGT 1226  
 Db 1078 GCTCTCTATCATCGCGCTTGTGCTGTGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1137  
 Qy 1227 ATCTCGTTTATTCGCGAAGAGAGGAGGCTGGAGTGCAGATGTGCTGACTGCTGCATTC 1286  
 Db 1138 GGACTCATTTACTGTGCGAAGAGAGGAGGCTGGAGTGCAGATGTGCTGCTGCATTCGCT 1197  
 Qy 1287 CTCACCTCAGGCTTCGAGCGCGCTCAGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCG 1346  
 Db 1198 CTCACCTCAGGCTTCGAGCGCGCTCAGCATCAAGCGCGCGCGCGCGCGCGCGCGCGCG 1257  
 Qy 1347 CTCACCATCAGCGCGAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1403  
 Db 1258 CTCACCATCAGCGCGAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1317  
 Qy 1404 CGGCGAGTGGCG 1463  
 Db 1318 AGGCGAGTGGAGCG 1377  
 Qy 1464 GGTGGCG 1523  
 Db 1378 GGGAGTGGCG 1437  
 Qy 1524 TCCCGCTCTCCACCGAGAACTACTTCGCTCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1583  
 Db 1438 TCCCGCTCTCCACCGAGAACTACTTCGCTCCCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 1497  
 Qy 1584 TATGGAGCTTCAACTTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1643  
 Db 1498 TACGGAGCTTCAACTTCTCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1557

QY 1644 CTCATCCCCCAGATCCCATACCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703  
Db 1558 CTCATACCCCGGATGCCATCCCGGAGGAAGATCTACAGATCTACCTCAGCTGCAC 1617  
QY 1704 AAGCCGGAAGAGTGTAGGTGGCCCTAGCTGGTGTGAGACCTGTGAGTCCCATCGTT 1763  
Db 1618 AAGCCGAAGAAGTGTAGGTGGCCCTAGCTGGTGTGAGACCTGTGAGTCCCATCGTT 1677  
QY 1764 AGCTGTGGACCCCTCGGCTCTGCTCTACCCGCCAGTCTATCCCTGCAATGGACCACTGT 1823  
Db 1678 AGCTGTGGCCCGCCAGAGTCTCTGCTACCCGCCAGTCTATCCCTGCAATGGACCACTGT 1737  
QY 1824 GGGAGCCCGAGCCCTGACAGCTGGAGCTCGGCTCAAAAAGCAGTCTGCGAGGCGAGC 1883  
Db 1738 GGAGAGCCCGAGCCCTGACAGCTGGAGTCTGGCCCTCAAAAAGCAGTCTGCGAGGCGAGT 1797  
QY 1884 TGGGAGCAGGATGTGTGACCTGGCGGAGGAGGCGCCCTCCACCTCTACTACTCCAG 1943  
Db 1798 TGGG---AGGATGTGTGACCTTGTGTGAGGAGTCACTTCCACCTCTACTACTCCAG 1854  
QY 1944 CTGGAGCCAGTCTGCTGCTGCTCTTACCGAGCAGCTGGCGGCTTGGCCCTGTGGGA 2003  
Db 1855 CTGGAGCCCGGCGCTGCTGCTCTTCTTACCGAGCAGCTGGCGGCTTGGCCCTGTGGGA 1914  
QY 2004 GAGGCCCTCAGCTGGCTGCCGCCAAGCGCCTCAAGCTGCTCTGTTTGGCGGCTGGCC 2063  
Db 1915 GAGGCCCTCAGCTGGCTGCCGCCAAGCGCCTCAGGCTCTTCTGTTTGGCTGGCC 1974  
QY 2064 TGCACCTCCCTCGAGTACAACATCGGGTCTACTGCTGCTGATGACACCCAGTGCATC 2123  
Db 1975 TGTAGCTCCCTTGAAGTACAACATCGAGTGTACTGCTTACAGACACCCAGCAGCTCTC 2034  
QY 2124 AAGAGGTGGTGCAGCTGGAGAGCAGCTGGGGGAGCAGCTGATCAGAGGACCGGTC 2183  
Db 2035 AAGAGGTGGTGCAGCTGGAGAGCAGCTGAGTGGACAGCTGATCAGAGGCTCGCGTC 2094  
QY 2184 CTGCATCTCAAGGACAGTTTACCAACCTGCGCCTATCCATCCAGATGTGCCAGCTCC 2243  
Db 2095 CTGCATCTCAAGGACAGTTTACCAACCTGCTCTCTCCATCCAGCTGCCAGCTCC 2154  
QY 2244 CTGTGGAAGAGTAACTCTTGTGAGTACAGGAGATCCCTTTTATCACTCTGGAAT 2303  
Db 2155 CTGTGGAAGAGCAAGTCTTGTGAGTACAGGAGATCCCTTTTATCACTCTGGAAT 2214  
QY 2304 GGCAGCAGCGTACTTGCAGTGCACCTTCAACCTGGAGCGTGTGAGCCCGAGCTAGT 2363  
Db 2215 GGCACCCAGAGTATCTGCAGTGCACCTTCAACCTGGAGCGGATCAACGCGCAGCAGC 2274  
QY 2364 GACCTGGCTGCAAGCTGTGGTGTGGCAGGTGGAGGCGACGGCAGAGCTTCAAGCATC 2423  
Db 2275 GACCTGGCTGCAAGGTGTGGTGTGGCAGGTGGAGGAGATGGCAGAGCTTCAAGCATC 2334  
QY 2424 AACTTCACATCACCAAGACACAAGTTTGTGAGCTGTGGTCTGGAGAGTGAAGCG 2483  
Db 2335 AACTTCACATCACTAAGGACACAAGTTTGTGAAATTTGTGGCTCTGGAGAGTGAAGCG 2394  
QY 2484 GGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTATTTCGGCAGAG 2543  
Db 2395 GGGGTCCAGCCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTATTTCGGCAGAG 2454  
QY 2544 ATAATTTCCAGCTGGACCCACCTCTGTAGGGGGGTGCGGAGTCTGCGCCAG 2603  
Db 2455 ATCATGGCAGTCTGGACCCACCTGTGAGCGGGGCGGAGTGGAGAACTCTAGCCAG 2514  
QY 2604 AAACTCACCTGGACAGCCATCTAGCTTCTTTTGTCTCAAGCCCGAGCCCAAGCCATG 2663  
Db 2515 AAACTCACCTGGACAGCCATCTTAGCTTCTTTTGTCTCAAGCCCGAGCCCTAGCCATG 2574  
QY 2664 ATCTCAACCTGTGGAGGCGGCGACTTCCCAACGCGCACTTCAAGCAGCTGGCTGCA 2723  
Db 2575 ATCTCAACCTGTGGAGGCGGCGACTTCCCAACGCGCACTTCCCAACGCGCACTGGCGAGC 2634

QY 2724 GCAGTGGCTGGATGGCGGCGAGCCAGAGCTGGGCTCTTTCACAGTGTGCGAGGCTGAGTGC 2783  
Db 2635 GCTGTGGCGGAGCTGGGCGCAACAGATGTGGGCTCTTTCACGCTGTGCGAGGCGGAGTGT 2694  
QY 2784 TGAGGCGGCGGCGAGGCGCG--ACACCTTACACTCTTACAGCTTTGGCACCCACCAAGACA 2841  
Db 2695 TGAGACCAAGCAGGCGCGGTAATGCTTACATCTTACAGCTTTGACACTGCCAGGACA 2754  
QY 2842 GGCAGAAAGCCGAGCAGGCGGCTTCCCAACACCGGGG 2878  
Db 2755 GGCAGAAACAGACAGGCGGCTTCCCAACACCGGGG 2791

RESULT 7  
US-10-240-154-15  
; Sequence 15, Application US/10240154  
; Publication No. US20030175741A1  
; GENERAL INFORMATION:  
; APPLICANT: Cochran et al.  
; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES  
; FILE REFERENCE: CKFW-P01-006  
; CURRENT APPLICATION NUMBER: US/10/240,154  
; CURRENT FILING DATE: 2001-04-02  
; PRIOR APPLICATION NUMBER: PCT/GB01/01486  
; PRIOR FILING DATE: 2001-04-02  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 15  
; LENGTH: 2697  
; TYPE: DNA  
; ORGANISM: Rattus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(2697)  
US-10-240-154-15

Query Match 76.0%; Score 2189.4; DB 15; Length 2697;  
Best Local Similarity 88.9%; Pred. No. 0;  
Matches 2403; Conservative 0; Mismatches 291; Indels 9; Gaps 3;  
QY 87 ATGGCGCTCCGCGCGGCTGTGGCAGCGCTCTGGGCAATAGTCTCGCCCTTGGCTC 146  
Db 1 ATGGCGCTCCGCGCGGCTGTGGCAGTGTCTTGGGCAATAGTCTCGCCCTTGGCTT 60  
QY 147 CGCGCTCGGTGGCCAGCAGAGTSCCACCCTGGCCAAACCCAGTGCCTGGTGCACACCG 206  
Db 61 CGTGTTCGGTGGCCAGCAGAGTSCCACCCTGGCCAAACCCAGTGCCTGGTGCACACCG 120  
QY 207 GACCTGCTTCCCGCTTCTCGTGGAGCCCGAGGATGTGTACATGTCAGAAAGCCA 266  
Db 121 GACCTGCTTCCCGCTTCTCGTGGAGCCCGAGGATGTGTACATGTCAGAAAGCCA 180  
QY 267 GTGCTGTGTGTGCAAGGCGGCTGCGCCCGCAGCAGATCTTCTTCAAGTGAACGGGAG 326  
Db 181 GTGTTGTGTGTGCAAGGCTGTGCTGCGCCAGCAGATCTTCTTCAAGTGAACGGGAG 240  
QY 327 TGGTGGCCAGGTGGACCAAGTGTGAGCGCAGCAGACGGGAGCAGTGTGTGAGCGG 386  
Db 241 TGGTGGCCAGGTGCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 300  
QY 387 ACCATGAGGTCCGCTTAAATGCTCAAGCAGCAGGTTCAGAGGATGTTCGGCTGGAG 446  
Db 301 ACCATGAGGTCCGCTTAAATGCTCAAGCAGCAGGTTCAGAGGATGTTCGGCTGGAG 360  
QY 447 GAACTACTGGTCCAGTGTGGCATGTGCTCGGCGCACCAAGAGTCAAGAGGCC 506  
Db 361 GAACTACTGGTCCAGTGTGGCATGTGCTCGGCGCACCAAGAGTCAAGAGGCC 420  
QY 507 TACATCCGATAGCCAGATTCGCAAGAACTTCGAGCAGGAGCGCTGGCCAGAGGTG 566  
Db 421 TACATCCGATAGCCAGATTCGCAAGAACTTCGAGCAGGAGCGCTGGCCAGAGGTG 480  
QY 567 TCCCTGGAGCAGGCGATCTGCTGCCCTGCGCTCCACGGAGGGCATCCCTTCAGCGGAG 626

Db 481 TCACGGAGCAGGCAATGTAACCTTGTGCCCCCAGAGGAATCCCCCAGCTGAG 540  
QY 627 GTGAGTGGCTCCCGAAACGAGACCTGTGTGACCGCTGCTGGACCCCAATGTATACATC 686  
Db 541 GTGAGTGGCTTCGAAATGAGACCTGTGTGACCGCTCCCTCGATCCCAATGTGTACATC 600  
QY 687 AGCGGGAGCAGCAGCTGTGTGTGCGACAGCGCCGCTTGTGTGACAGCGCCAACTACACC 746  
Db 601 AGCGGGAGCAGCAGCTGTGTGTGCGACAGCGCCGCTTGTGTGACAGCGCCAACTACACC 660  
QY 747 TGCCTGGCCAGAAACATGTGTGACAGCTGTGCGACAGCGCTCCCTGTGTATCTGTCTAC 806  
Db 661 TGTGTGGCCAGAAACATGTGTGACAGCTGTGCGACAGCGCTCCCTGTGTATCTGTCTAT 720  
QY 807 GTGAACGGTGGTGTGACGTGTGACAGCGCTGTGCGACAGCGCCCACTGTGTGGGCGC 866  
Db 721 GTGAACGGTGGTGTGACGTGTGACAGCGCTGTGCGACAGCGCTCCCTGTGTGGGCGT 780  
QY 867 GGCTGGCAGAAACGAGCGGAGCTGTGACCAACCGCGGCGCTTCAACGGGGCGCTTTC 926  
Db 781 GGCTGGCAGAAACGAGCGGAGCTGTGACCAACCGCGGCGCTTCAACGGGGCGGCTTTC 840  
QY 927 TGTGTGGGCGAGAAATGTCATGACCGCACCGTCTCTCTCTGCTGTCTCTGTGTGACGCGC 986  
Db 841 TGTGTGGGCGAGAAATGTCAGAAACAGCTGTGCGCACCTCTG---TGCCCACTGTGATGGG 897  
QY 987 AGCTGGAGCCGCTGGAGCAAGTGTGCGCTGTGTGGCTGTGACCTGACACCCACTGGGGAGC 1046  
Db 898 AGCTGGAGTTCGTGGAGTAAAGTGTGAGCTGTGTGGCTTGTGACTGTGACACCCACTGGGGAGC 957  
QY 1047 CGTGTGTGTGACCCAGACCCCGCAACGAGGAGGAGGAGTGCAGGGCACTGACCTG 1106  
Db 958 CGCGAGTGTCTGACCCAGCACCCCGCAATGAGGTGAGGAGTGTGCGGGTGTGACCTG 1017  
QY 1107 GACACCCGCAACTGTACCACTGACCTCTGTGTACAGTGTCTTCTGCGCTGAGGACGTG 1166  
Db 1018 GACACCCGCAACTGTACCACTGACCTCTGTGTGACAGTGTCTTCTGCGCTGAGGAGCTG 1077  
QY 1167 GCGCTCTATGTGGGCTCTATCGCCGTGTGCGCTGTGCTGTGCTGTGTGTGTGTGTGTGT 1137  
Db 1078 GCTCTCTACATGGCCTTGTGCTGTGTGCTGTGTGCTCTTCTTGTGTGTGTGTGTGTGT 1137  
QY 1227 ATCTCGTTTATGCGGAGAGAGGAGGCTGTGACTGTGATGTGTGCTGACTGTCTGCTCAT 1286  
Db 1138 GGACTCAITTTACTGTGCAAGAGAGGAGGCTGTGACTTCGATGTGCGCCGACTGTCTCATC 1197  
QY 1287 CTCACCTCAGGCTTCCAGCCCGTGTGATCAAGCCCGAGCAAGCAGCAACCCCACTGTG 1346  
Db 1198 CTCACCTCGGCTTCCAGCTGTGAGATCAAGCCCGAGCAAGCAGCAACCCCACTGTG 1257  
QY 1347 CTCACCATCAGCCGAGCCTCAGCACAC---ACCACTTACAGGGGAGTCTCTGTCTCC 1403  
Db 1258 CTCACCATCAGCCAGACCTCAGCACACCACTTACCACTTACCACTTACCACTTACCACT 1317  
QY 1404 CGGAGAGTGGGCGGAGCCCGAGCTTCCAGCTCAGCTCAGCTTCCAGCTTCCAGCTTCCAG 1463  
Db 1318 AGCAGAGTGGAGCCCGAGCCCAAGTTCAGCTCTCTAATGTGTCACTGTCTCAGCCCACTG 1377  
QY 1464 GGTGGGCGGCGCACACACTGTGACCACTGTCTCCCACTTCTGAGCGGAGGAGTGTCTG 1523  
Db 1378 GGGAGTGGGCGGCACTGTGTGACCACTGTGACCACTTACCACTTACCACTTACCACT 1437  
QY 1524 TCCCGCTCTCCACCCAGACACTATCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1583  
Db 1438 TCCCGCTCTCCACCCAGACACTATCTTCCGCTTCCGCTTCCGCTTCCGCTTCCGCTTCCG 1497  
QY 1584 TATGGAGCTTCAACTTCTCGGGGCGGCTGTATGATCTTCAATCACTATCACTATCAGCTC 1643  
Db 1498 TAGGGAGCTTCAACTTCTCGGGGCGGCTGTATGATCTTCAATCACTATCACTATCAGCTC 1557  
QY 1644 CTCATCCCCCAGATGCATACCCCGAGGAGAGATCTATGAGATCTACCTTCACTGCTGCAC 1703

1558 CTCATACCCCGGATGCCATCCCCGAGGAAAGATCTACGAGATCTACCTCACACTGCAC 1617  
1704 AAGCGGAGAGCTGAGGTTGCCCTAGCTGCTGTGAGCCCTGCTGAGCCCTGCTGAGTCCCATCGTT 1763  
1618 AAGCCAGAGAGCTGAGGTTGCCCTAGCTGCTGTGAGCCCTGCTGAGTCCCATCGTT 1677  
1764 AGCTGTGACCCCTGGCGCTCTGTCTACCCGGGCGAGTCACTCTGTGGCTATGAGCACTGT 1823  
1678 AGCTGTGGGCCCCAGGAGTCTGTCTACCCGGGCGAGTCACTCTTGTCAATGAGCACTGT 1737  
1824 GGGAGCCGAGCCTGTGACAGCTGTGAGCCTGTGCGCTCAAAAAGCAGTGTGTGAGGAGCAG 1883  
1738 GAGAGCCGAGCCTGTGACAGCTGTGAGTCTGTGCGCTCAAAAAGCAGTGTGTGAGGAGCAGT 1797  
1884 TGGAGAGAGGATGTCTGACCTGTGCGGAGGAGGCGCCCTCCACTCTACTACTGTGCCAG 1943  
1798 TGGG---AGGAGTGTCTGACCTTGTGAGGAGTCACTTCCCACTCTACTACTGTGCCAG 1854  
1944 CTGAGGCGCAGTCCCTGTCTTACCTTTCACCGAGCAGCTGGGCGCTTTGCCCTGTGGGA 2003  
1855 CTGAGGCGCGGGCTGTCTATGTCTTACCGAGCAGCTGGGCGCTTTGCCCTGTGGGA 1914  
2004 GAGGCCCTGAGCTGGCTGCGGCCCAAGCGCTCAAGCTGTCTTCTGTGTGCGCGCTGCGC 2063  
1915 GAGGCCCTGAGCTGGCTGCGGCCCAAGCGCTCAAGCGCTCTTCTGTGTGCTCCCTGGCC 1974  
2064 TGCACCTCCCTGAGTACAACTCCGGGTCTACTGCTGATGACACCCAGATGCACATC 2123  
1975 TGTGCTCCCTTGTGTACAACTCCGAGTGTACTGCTTACAGCAGCACCACAGAGCTCTC 2034  
2124 AAGAGGCTGTGAGTGTGAGGAGCAGCTGGGCGGAGCAGCTGATCCAGGAGCCACGGGTC 2183  
2035 AAGAGGCTGTGAGTGTGAGGAGCAGCTGAGTGTGAGCAGCTGATCCAGGAGCCTCGCTC 2094  
2184 GTGCACTTCAAGGACAGTTACCAACCTGTGCGCTTATCCATCCAGATGTGCCAGCTCC 2243  
2095 GTGCACTTCAAGGAGAGTTACCAACCTGTGCTCTCTCCATCCAGAGCTGCGCAGCTCC 2154  
2244 CTGTGGAGAGTAACTCTCTGCTGACCTACAGGAGATCCCTTTTATCACATCTCGAAT 2303  
2155 CTGTGGAGAGCAGTCTTGTGCTGAGTACAGAGATCCCTTTTATCACATCTGGAAC 2214  
2304 GGCAGCAGCGGTACTTGTGACTGTGACCTTACCTTGTGAGCTGTGTGAGCAGCTAGT 2363  
2215 GGCACCCAGCAGTATCTGCACTGTGCTTACCTTGTGAGCGCATCAACGCCAGCAGC 2274  
2364 GACCTGGCTGTCAAGCTGTGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2423  
2275 GACCTGGCTGTCAAGCTGTGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2334  
2424 AACTTCAACATCACCAAGGAGACAAAGTGTGTGCTGTGAGCTGTGAGAGTGAAGCG 2483  
2335 AACTTCAACATCACCAAGGAGACAAAGTGTGTGAGTGTGTGAGTGTGAGAGTGAAGGG 2394  
2484 GGGGTCCCAGCCTGTGTGGGCCCCAGTGTCTTCAAGATCCCTTCTCTAATGTGGGAGAG 2543  
2395 GGGGTCCCAGCCTGTGTGGGCCCCAGTGTCTTCAAGATCCCTTCTCTAATGTGGGAGAG 2454  
2544 ATAAATTCAGCTGTGAGCCCACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2603  
2455 ATCATCGCAGTGTGAGCCCACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2514  
2604 AACTTCACTGTGAGCAGCATCTGAGCTTCTTGTGCTTCAAGCCAGCCAGCCAGCCAG 2663  
2515 AACTTCACTGTGAGCAGCATCTGAGCTTCTTGTGCTTCAAGCCAGCCAGCCAGCCAG 2574  
2664 ATCTTCAACTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2723  
2575 ATCTTCAACTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2634  
2724 GAGTGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2783  
2635 GCTGTGGGCGGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2694



Qy	2430	AAATACCAAGAGACAAAGTTTCTGAGACTGCTCTGAGAGTGAAGCGGGTC	2489
Db	1312	AAATCACCAGGACACAAGGTTTGTCTGAGCTGTCTGGCTCTCGAGAGTGAAGCGGGTC	1371
Qy	2490	CCAGCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTTCCGACAGAAATAATT	2549
Db	1372	CCAGCCTTGGTGGGCCCAAGTGCCCTTCAAGATCCCTTCTCATTTCCGACAGAGTAATT	1431
Qy	2550	TCCAGCCTGGACCCACACCTCTAGGCGGGGTGCCGACTGGCGGACTCTGGGCCAGAAACTC	2609
Db	1432	TCCAGCCTGGACCCACACCTCTAGGCGGGGTGCCGACTGGCGGACTCTGGGCCAGAAACTC	1491
Qy	2610	CACCTGGACAGCAATCTCAGCTTCTTTTGCCTTCCAAGCCCGAGCCCAAGACCATGATCCTC	2669
Db	1492	CACCTGGACAGCCATCTCAGCTTCTTTTGCCTTCCAAGCCCGAGCCCAAGACCATGATCCTC	1551
Qy	2670	AACCTGTGGAGGCGCGGCACCTTCCCCAAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	2729
Db	1552	AACCTGTGGAGGCGCGGCACCTTCCCCAAACGGCAACCTCAGCCAGCTGGCTGCAGCAGTG	1611
Qy	2730	GCTGGACTGGCCAGCCAGACGCTGGCCCTC-TTCACAGTG-TCCGAGGCTGAGTGCTGAG	2787
Db	1612	GCTGGGACTGGCCAGCAGGACGGTGGCTTCTTTTCAACAGTGTTCCGAGGCTGAGTGCTGAG	1671
Qy	2788	GCCGGCCAGGCCCGACACCTACACTCTCACCAGCTTTTGGCA--CCCAACAGGACAGGCCA	2845
Db	1672	GCCGGCCAGG--CGAACNCTACAAATTTTACAGTTTTTGGGAACCCACCAAGGACAGGCCA	1729
Qy	2846	GAAGCCGAGACAGGGGCCCTTCCCCACACCGGGGAGA	2881
Db	1730	GAAGCCGAGACAGGGGCTTTTCCCAAAAACCGGGGAGA	1765

## RESULT 9

US-10-256-702-2

: Sequence 2, Application US/10256702

: Publication No. US20030059859A1

GENERAL INFORMATION:

GENERAL INFORMATION: APPLICANT: Tessier-Lavigne Marc

APPLICANT: Tessier-Lavigne, Marc  
 Tessier-Lavigne, David

Leonardo, E. David

Hink, Lindsay

Masu, Masayuki

Kazuko, Keino-Masu

TITLE OF INVENTION: Netrin Receptors

NUMBER OF SEQUENCES: 8

NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:

ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP

STREET: 268 BUSH STREET, SUITE 3200

CITY: SAN FRANCISCO

STATE: CALIFORNIA

COUNTRY: USA

COUNT: 000  
ZTP: 94104; ;  
ZIP: 94104  
COMPUTER READABLE FORM.

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5 1/4 INCH DISK

```

;
;
MEDIUM TYPE: Floppy disk

```

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/256,702

AFFIDAVIT NUMBER: 05/10/2002  
FILING DATE: 27-Sep-2002

FILING DATE: 27-sep-2002  
CLASSIFICATION: ~~Unknown~~

```

;
;
CLASSIFICATION: <UNKNOWN>
REF ID: A66763

```

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US/09/933,261

FILING DATE: 20-Aug-2001

APPLICATION NUMBER: 08/808,982

FTLTING DATE: <Unknown>

FILING DATE: 04/01/2015  
ATTORNEY/AGENT INFORMATION:

ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN RICHARD A

NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 26 007

```

; INFORMATION FOR SEQ ID NO: 2:
;-----
; SEQUENCE CHARACTERISTICS:
;     LENGTH: 1787 base pairs
;     TYPE: nucleic acid
;     STRANDEDNESS: double
;     TOPOLOGY: linear
;-----
; MOLECULE TYPE: cdna
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
;
US-10-256-703-2

```

Query Match	54.5%;	Score 1570.4;	DB 15;	Length 1787;
Best Local Similarity	96.3%;	Pred. No. 0;		
Matches 1721;	Conservative	0;	Mismatches	36;
			Indels	19;
			Gaps	11;

QY	1114	GCACCTGTACCACTGACCTCTGTTGATACACAGTCTCTTCTGCGCCTCTAGAGGACGTGGCCCTCT	1171
DB	1	GCACCTGTACCACTGACCTCTG-TGACACTGCTTCTTGCCCTCTGAGACGTGGCCCTCT	59
QY	1174	ATGTGGGCTCATATCGCCGTGGCCGTCTGCTGTCTGTGTGTCTGTCTCTCATCTCTCG	1233
DB	60	ATGTGGGCTCATATCGCCGTGGCCGTCTGCTGTCTGTGTCTGTCTCTCATCTCTCG	119
QY	1234	TTTATTGCCGAGAGAGAGGGGCTTGGACTCAGATCTGGCTGACTCGTCCATCTCACCT	1293
DB	120	TTTATTGCCGAGAGAGAGGGGCTTGGACTCAGATGTGGCTGACTCGTCCATCTCACCT	179
QY	1294	CAGGCTTCCAGGCCGTCTCAGCATC-AAGCCCGAGCAAGACAGACAAACCCCATCTGCTCAC	1352
DB	180	CAGGCTTCCAGGCCGTCTCAGCATCTAAGCCCGAGCAAGACAGACAAACCCCATCTGCTCAC	239
QY	1353	ATCCAGCCGGAACCTCAG---CACACACACACTACAGGGGAGTCTCTGTCCCGGCAG	1409
DB	240	ATCCAGCCGGAACCTCAGCAGCAGCACCACACTTACAGGGGAGTCTCTGTCCCGGCAG	299
QY	1410	GATGGGCCAGCCCCAAGTTCCAGCTCACCAAATGGGCACCTGTCTCAGCCCCCTGGGTGC	1469
DB	300	GATGGGCCAGCCCCAAGTTCTCAGCTCACCAAATGGGCACCTGTCTCAGCCCCCTGGGTGC	359
QY	1470	GGCCGCCACACACTGCACACAGCTCTCCACCTCTGAGGCCGAGGAGTTGCTCTCCGC	1529
DB	360	GGCCGCCACACACTGCACACAGCTCTCCACCTCTGAGGCCGAGGAGTTGCTCTCCGC	419
QY	1530	CTCTCCACCCAGAACTACTTCCGCTCCTCGCCCGGAGCACCAACATGACCTATGGG	1589
DB	420	CTCTCCACCCAGAACTACTTCCGCTCCTCGCCCGGAGCACCAACATGACCTATGGG	479
QY	1590	ACCTTCAACTTCTTCGGGGCCGCTCATGATCCTTAATACAGTATCAGCTCTCTCATC	1649
DB	480	ACCTTCAACTTCTTCGGGGCCGCTCATGATCCTTAATACAGTATCAGCTCTCTCATC	539
QY	1650	CCCCCAGATGCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTTGCAACAGCG	1709
DB	540	CCCCCAGATGCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTTGCAACAGCG	599
QY	1710	GAAAGCTGAGTTGCCCTAGCTGGTGCTCAGACCTGTGAGTCCCATCTGTTAGTGT	1769
DB	600	GAAAGCTGAGTTGCCCTAGCTGGTGCTCAGACCTGTGAGTCCCATCTGTTAGTGT	659
QY	1770	GGACCCCTGGCGTCTCTCTCACCGGCCAGTCACTCTGGCTATGGACCACTGTGGGAG	1829
DB	660	GGACCCCTGGCGTCTCTCTCACCGGCCAGTCACTCTGGCTATGGACCACTGTGGGAG	719
QY	1830	CCGAGCCTTGACAGCTGGAGCTTGGCTCAAAAGCAGTCTGTCGAGGCGACGTGGAG	1889
DB	720	CCGAGCCTTGACAGCTGGAGCTTGGCTCAAAAGCAGTCTGTCGAGGCGACGTGGAG	778
QY	1890	CAGATGTGCTGCACTTGGGAGAGAGCGCCCTCCACCTCTACTCTGCCAGCTGGAG	1949
DB	779	GATG-----TCTGCACTTGGGAGAGAGGCGCCCTCCACCTCTACTCTGCCAGCTGGAG	834
QY	1950	GCCAGTGCCTCTACGTTCTTCCAGCAGCAGCTGGGCGCTTTGGCTGGGAGAGGCC	2009
DB	835	GCCAGTGCCTCTACGTTCTTCCAGCAGCAGCTGGGCGCTTTGGCTGGGAGAGGCC	894



QY 2010 CTGAGCTGGCTGCGCCAGCGCCTCAAGCTGCTTCTGTTTGGCGGGTGGCTGCACC 2069  
Db 895 CTGAGCTGGCTGCGCCAGCGCCTCAAGCTGCTTCTGTTTGGCGGGTGGCTGCACC 954  
QY 2070 TCCTCGAGTACAACTCCGGGTCTAGCTGCTGATGACACCCAGATGCACTCAAGGAG 2129  
Db 955 TCCTCGAGTACAACTCCGGGTCTAGCTGCTGATGACACCCAGATGCACTCAAGGAG 1014  
QY 2130 GTGGTGCAGCTGAGAGCAGCTGGGGGACAGCTGATCCAGAGGACAGGGTCTGGCAC 2189  
Db 1015 GTGGTGCAGCTGAGAGCAGCTGGGGGACAGCTGATCCAGAGGACAGGGTCTGGCAC 1074  
QY 2190 TTCAAGGACAGTTTACCAACCTCGCCCTATCCATCCAGATGTGCCAGCTCCCTGTGG 2249  
Db 1075 TT-AAGGACAGTTTACCAACCT--GCCCTATCATCCAGATGTGCCAGCTCCCTGTGG 1131  
QY 2250 AAGAGTAAGCTCTGTGCTACCAAGGAGATCCCTTTTATCACATCTGGAATGGCAAG 2309  
Db 1132 AAGAGTAAGCTCTGTGCTACCAAGGAGATCCCTTTTATCACATCTGGAATGGCAAG 1191  
QY 2310 CAGCGTACTTGCAGTGCACCTTACCCCTGGAGCGTGTGAGCCCGAGCTAGTGAAGTGG 2369  
Db 1192 CAGCGTACTTGCAGTGCACCTTACCCCTGGAGCGTGTGAGCCCGAGCTAGTGAAGTGG 1251  
QY 2370 GCTGCAAGCTGTGGGTGTGGAGTGTGGAGGTGGAGGCGACGGGACAGCTTCAAGCTTCA 2429  
Db 1252 GCTGCAAGCTGTGGGTGTGGAGTGTGGAGGTGGAGGCGACGGGACAGCTTCAAGCTTCA 1311  
QY 2430 AACATCACCAAGGACAAAGGTTTGTGAGTGTGGTCTGTGAGAGTGAAGCGGGGGTTC 2489  
Db 1312 AACATCACCAAGGACAAAGGTTTGTGAGTGTGGTCTGTGAGAGTGAAGCGGGGGTTC 1371  
QY 2490 CCAGCCCTGTGGGCGGCGGCTTCAAGATCCCTTCTCATCTTCGGCAGAGATAATT 2549  
Db 1372 CCAGCCCTGTGGGCGGCGGCTTCAAGATCCCTTCTCATCTTCGGCAGAGATAATT 1431  
QY 2550 TCAGCTGTGAGCCACCTGTAGGGGGGTGGCGACTGGCGGACTTGGCCCGAGAACTTC 2609  
Db 1432 TCAGCTGTGAGCCACCTGTAGGGGGGTGGCGACTGGCGGACTTGGCCCGAGAACTTC 1491  
QY 2610 CACTGTGAGCAGCATCTCAGCTTCTTTGGCTTCAAGATCCCTTCTCATCTTCGGCAGAG 2669  
Db 1492 CACTGTGAGCAGCATCTCAGCTTCTTTGGCTTCAAGATCCCTTCTCATCTTCGGCAGAG 1551  
QY 2670 AACCTGTGGAGCGCGGCACTTCCCAAGCGCACTTCCCAAGCGGCGGCGGCGGCGG 2729  
Db 1552 AACCTGTGGAGCGCGGCACTTCCCAAGCGCACTTCCCAAGCGGCGGCGGCGGCGG 1611  
QY 2730 GCTGAGCTGGGCGAGCAGCTGGCCCTC-TTCACAGTG-TCCGAGGTGAGTGTCTGAG 2787  
Db 1612 GCTGAGCTGGGCGAGCAGCTGGCCCTC-TTCACAGTG-TTCGAGGTGAGTGTCTGAG 1671  
QY 2788 GCCGCGAGGCGGCGGCACTTCCCAAGCGCACTTCCCAAGCGGCGGCGGCGGCGG 2845  
Db 1672 GCCGCGAGG--CGAACAATAATTTTACAGTTTGGGAACCCCAAGGAGCAGGCA 1729  
QY 2846 GAAGCGGAGAGGCGGCGGCTTCCCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 2881  
Db 1730 GAAGCGGAGAGGCGGCTTCCCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1765

## RESULT 10

US-10-296-115-365  
; Sequence 365, Application US/10296115  
; Publication No. US20040053248A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq Inc  
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides  
; FILE REFERENCE: 784PCT  
; CURRENT APPLICATION NUMBER: US/10/296,115  
; CURRENT FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: US09/488,725

; PRIOR FILING DATE: 2000-01-21  
; PRIOR APPLICATION NUMBER: US09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; NUMBER OF SEQ ID NOS: 1478  
; SEQ ID NO 365  
; LENGTH: 1321  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-296-115-365

Query Match 41.4%; Score 1193.6; DB 13; Length 1321;  
Best Local Similarity 97.8%; Pred. No. 2.6e-287;  
Matches 1295; Conservative 0; Mismatches 19; Indels 10; Gaps 8;  
QY 1476 CACACACTGCACACAGCTTCTCCACTCTGAGCGCGAGAGTTCGTCTCCCGCTCTCC 1535  
Db 1 CACACACTGCACACAGCTTCTCCACTCTGAGCGCGAGAGTTCGTCTCCCGCTCTCC 60  
QY 1536 ACCCAGAACTACTTCGGCTCCCTGCCCGGAGGACACCAACATGACCTATGGGACTTC 1595  
Db 61 ACCCAGAACTACTTCGGCTCCCTGCCCGGAGGACACCAACATGACCTATGGGACTTC 120  
QY 1596 AACTTCTCTCGGGGCGCGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCCCCCA 1655  
Db 121 AACTTCTCTCGGGGCGCGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCCCCCA 180  
QY 1656 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCGGAGAC 1715  
Db 181 GATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCACGCTGCACAAGCGGAGAC 240  
QY 1716 GTGAGGTTGCCCTAGCTGGCTGTGAGCCCTGTGAGTCCCATCTTACGTGTGAGCC 1775  
Db 241 GTGAGGTTGCCCTAGCTGGCTGTGAGCCCTGTGAGTCCCATCTTACGTGTGAGCC 300  
QY 1776 CCT-GCGCTCTCTGCTCACCGGCGAGTCTATCCT-GGCTATGACCACTGT-GGGGAGCC 1832  
Db 301 CCTGGGCTCTGCTTACCGGCGAGTCTATCTGGGTATGACCACTGTGGGAGCC 360  
QY 1833 AGCCTTGACAGCT-GGAGCTCTGCGCTCAAAAGCAGTCTGCGAGGGCAGCTGGAGCA 1891  
Db 361 AGCCTTGACAGCTGCGGAGCTGCGCTCAAAAGCAGTCTGCGAGGGCAGCTGGG--A 417  
QY 1892 GGATGTGTGACCTTGGGCGAGGCGGCTTCCCACTCTACTACTGCGAGCTGGAGC 1951  
Db 418 GGATGTGTGACCTTGGGCGAGGCGGCTTCCCACTCTACTACTGCGAGCTGGAGC 477  
QY 1952 CAGTCCCTGTGAGTCTTTCACCGAGCAGTGGGCGCTTTCCTGCTGGGAGAGCCCT 2011  
Db 478 CAGTCCCTGTGAGTCTTTCACCGAGCAGTGGGCGCTTTCCTGCTGGGAGAGCCCT 537  
QY 2012 CAGCTGTGCTCGCGCAAGCGCTCAAGCTCTTCTGTTTGGCGCGTGGCTGCACCTC 2071  
Db 538 CAGCTGTGCTCGCGCAAGCGCTCAAGCTCTTCTGTTTGGCGCGTGGCTGCACCTC 597  
QY 2072 CCTCGAGTACAACTCCGGTCTACTGCTGATGACACCCAGATGCACTCAAGAGGT 2131  
Db 598 CCTCGAGTACAACTACTGGTCTACTGCTGATGACACTCAGATGCACTCAAGAGT 657  
QY 2132 GGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCAGGGTCTGCACTT 2191  
Db 658 GGTGAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGGAGCAGGGTCTGCACTT 717  
QY 2192 CAAGGACAGTTTACCAACCTTGGCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGAA 2251  
Db 718 CAAGGACAGTTTACCAACCTTGGCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGAA 777  
QY 2252 GAGTAAGCTCTTGTGAGTACCAAGAGATCCCTTTTATCACTATGGAATGGACGCA 2311  
Db 778 GAGTAAGCTCTTGTGAGTACCAAGAGATCCCTTTTATCACTATGGAATGGACGCA 837  
QY 2312 GCGGTACTTGCAGTGCACCTTACCCCTGGAGCGTGTGAGCCCGAGCACTAGTGAAGTGGC 2371  
Db 838 GCGGTACTTGCAGTGCACCTTACCCCTGGAGCGTGTGAGCCCGAGCACTAGTGAAGTGGC 897

QY 2372 CTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGACGGGAGAGCTTTCAGCATCAACTTCAA 2431  
DB 898 CTGCAAGCTGTGGGTGTGGAGGTGGAGGGCGACGGGAGAGCTTTCAGCATCAACTTCAA 957  
QY 2432 CATCAACAGGACACAAGGTTTGTGAGCTGTGCTGTGGAGAGTGAAGCGGGGTCCC 2491  
DB 958 CATCAACAGGACACAAGGTTTGTGAGCTGTGCTGTGGAGAGTGAAGCGGGGTCCC 1017  
QY 2492 AGCCCTGTGGGCCCCAGTGTCTCAAGATCCCTTCTCTCATTTGGCGAAGATAATTTC 2551  
DB 1018 AGCCCTGTGGGCCCCAGTGTCTCAAGATCCCTTCTCTCATTTGGCGAAGATAATTTC 1077  
QY 2552 CAGCTGAGACCAACCTGTAGGCGGGGTGCGAGCTGTGGGAGTCTGGGCCAGAACTCCA 2611  
DB 1078 CAGCTGAGACCAACCTGTAGGCGGGGTGCGAGCTGTGGGAGTCTGGGCCAGAACTCCA 1137  
QY 2612 CCTGACAGCAATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCCATGATCTCAA 2671  
DB 1138 CCTGACAGCAATCTCAGCTTCTTTGCTTCCAAAGCCAGCCGCCACAGCCATGATCTCAA 1197  
QY 2672 CCTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCGAGAGTGGC 2731  
DB 1198 CCTGTGGAGGCGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGCTGCGAGAGTGGC 1257  
QY 2732 T-GGACTGGGCGAGCAGAGCTGGCTC-TTCACAGTG-TGGAGGCTGAGTGTGAGG 2788  
DB 1258 TGGGACTGGGCGAGCAGAGCTGGCTTCTTTTACAGTGTTCGGAGGCTGAGTGTGAGG 1317  
QY 2789 CCGG 2792  
DB 1318 CCGG 1321

RESULT 11  
US-09-833-381-1806  
; Sequence 1806, Application US/09833381  
; Patent No. US2002013209A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US2002013209A1e1 Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-1119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1806  
; LENGTH: 1282  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(1282)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1806

Query Match 39.8%; Score 1147.8; DB 9; Length 1282;  
Best Local Similarity 97.7%; Pred. No. 6.7e-276;  
Matches 1259; Conservative 0; Mismatches 18; Indels 11; Gaps 9;

QY 1510 CCGAGAGTTCGTCTCCGCTTCCACCCAGAACTTCCGCTCCCTCCCGGAGGCA 1569  
DB 1 CCGAGAGTTCGTCTCCGCTTCCACCCAGAACTTCCGCTCCCTCCCGGAGGCA 60  
QY 1570 CCAGCAACATGACTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATA 1629  
DB 61 CCAGCAACATGACTATGGACCTTCAACTTCTCGGGGCGGCTGATGATCCCTAATA 120  
QY 1630 CAGGTATCAGCTCTCTATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCT 1689  
DB 121 CAGGAATCAGCTCTCTATCCCCCAGATGCCATACCCCGAGGAAGATCTATGAGATCT 180

QY 1690 ACCTCAGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTGAGACCTGEC 1749  
DB 181 ACCTCAGCTGCACAAGCCGGAAGACGTGAGGTTGCCCTAGCTGGCTGTGAGACCTGEC 240  
QY 1750 TGAGTCCCATCGTTAGCTGTGAGACCCCTT-GGCGTCTGTCTACCCGGCGAGTATCTCT- 1807  
DB 241 TGAGTCCCATCGTTAGCTGTGAGACCCCTTGGCGCTGTCTGCTCACCCGCCAGTATCTCTG 300  
QY 1808 GGCTATGACCACTGTGGGGAGCCAGCCCTGACAGCTT-GGAGCCTGGCGCTCAAAAAGC 1866  
DB 301 GGCTATGACCACTGTGGGGAGCCAGCCCTGACAGCTGTGGAGCTTGGCCCTCAAAAAGC 360  
QY 1867 AGTCTGTGAGGCGAGCTGGGAGAGGATGTGCTGCACCTGTGGCGAGGAGCGCCCTTCC 1926  
DB 361 AGTCTGTGAGGCGAGCTGGGAGGATGTGCT--TGACCTGGCGAGGAGCGCCCTTCC 418  
QY 1927 ACCTCTACTTGTGCGAGCTGGAGGCGAGTGCCTGTCTAGTCTTTCACCGAGAGCTGGGC 1986  
DB 419 ACCTCTACTTGTGCGAGCTGGAGGCGAGTGCCTGTCTAGTCTTTCACCGAGAGCTGAGCC 478  
QY 1987 GCTTTGCCCTGTGGGAGAGCCCTCAGCTGGCTGCCCGCAAGCGCTCAAGCTGCTTC 2046  
DB 479 GCTATGSCCTGTGGGAGAGGCGCTCAGCTGGCTGCCCGCAAGCGCTCAAGCTGCTTC 538  
QY 2047 TGTTTGGCGCGGTGGCTGCACCTCCCTCGAGTACAACAATCCGGGTCTACTGCTGCATG 2106  
DB 539 TGTTTGGCGCGGTGGCTGCACCTCCCTCGAGTACAACAATCTGCTTACTGCTGCATG 598  
QY 2107 ACACCCAGATGACTCAAGAGGTGTGCTGAGCTGGAGAGAGCTGGGGGAGACCTGA 2166  
DB 599 ACACCTCAGATGACTCAAGCTGTGCTGAGTGTGCTGAGAGAGCTGGGGGAGACCTGA 658  
QY 2167 TCCAGAGCCAGGGTCTGCACTTTCAGGAGAGTACCAACAACCTGCCCTATTCATCC 2226  
DB 659 TCCAGAGCCAGGGTCTGCACTTTCAGGAGAGTACCAACAACCTGCCCTATTCATCC 718  
QY 2227 AGGATGTGCCAGCTCCCTGTGGAAGAGTAACTCCTTGTGCTGAGTACCAAGAGATCCCT 2286  
DB 719 AGGATGTGCCAGCTCCCTGTGGAAGAGTAACTCCTTGTGCTGAGTACCAAGAGATCCCT 778  
QY 2287 TTTATCACAATGTGGAATGCAAGCGGTACTTTGCACTGCACTTTCAGCTACCAAGAGATCCCT 2346  
DB 779 TTTATCACAATGTGGAATGCAAGCGGTACTTTGCACTGCACTTTCAGCTACCAAGAGATCCCT 838  
QY 2347 TCAGCCCCAGCACTAGTGTGCTGCAAGCTGTGGGTGTGGAGGTGGAGGGGCAAG 2406  
DB 839 TCAGCCCCAGCACTAGTGTGCTGCAAGCTGTGGGTGTGGAGGTGGAGGGGCAAG 898  
QY 2407 GGCAGAGCTTTCAGCATCAACTTCAACATCAACAAGGACACAAGTTTGTGAGCTGTGG 2466  
DB 899 GGCAGAGCTTTCAGCATCAACTTCAACATCAACAAGGACACAAGTTTGTGAGCTGTGG 958  
QY 2467 CTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGGCCCGAGTGTTCAGATCCCT 2526  
DB 959 CTCTGAGAGTGAAGCGGGGTCCAGCCCTGTGGGGCCCGAGTGTTCAGATCCCT 1018  
QY 2527 TCCTCATTCGGAGAGATAA-TTTCAGCTGCAACCCCTGTAGGGGGGTGGCGAC 2585  
DB 1019 TCCTCATTCGGAGAGATAA-TTTCAGCTGCAACCCCTGTAGGGGGGTGGCGAC 1078  
QY 2586 TGGCGG-ACCTGTGGCCCCAGAAATCCACCTGGAGAGCCATCTCAGCTTCTTTTGGCTCCAA 2644  
DB 1079 TGGCGGAATCTTGGCCCCAGAAATCCACCTGGAGAGCCATCTCAGCTTCTTTTGGCTCCAA 1138  
QY 2645 GCCAGCCCCAGACCATGATCTCAACTGTGGGAGCGCGCACTTCCCCAACGGCAA 2704  
DB 1139 GCCAGCCCCAGACCATGATCTCAACTGTGGGAGCGCGCACTTCCCCAACGGCAA 1198  
QY 2705 CCTCAGCAGCTGGCTGCAAGCTGGCTGGGCGAGCCAGCAGAGCTGGCTTCTTTCAC 2764  
DB 1199 CCTCAGCAGCTGG-TGCAGCAGTGGCTGGACTTGGG--AGCAGAGCTGG-CTCTTTCAC 1254

QY 2765 AGTGTGGAGGCTGAGTGTGAGGCGG 2792  
Db 1255 AGTGTGGAGGCTGAGTGTGAGGCGG 1282

## RESULT 12

US-10-087-684-1  
; Sequence 1, Application US/10087684  
; Publication No. US20040029116A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David J.  
; APPLICANT: Grosse, William M.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Casman, Stacie, J.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gangolli, Esha A.  
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-214 CIP  
; CURRENT APPLICATION NUMBER: US/10/087,684  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: 60/253,834  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: 60/250,926  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/264,180  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/313,656  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: 60/327,456  
; PRIOR FILING DATE: 2001-10-05  
; NUMBER OF SEQ ID NOS: 220  
; SOFTWARE: CuraSeqList version 0.1  
; SEQ ID NO 1  
; LENGTH: 2860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (59)..(2857)  
US-10-087-684-1

Query Match 30.9%; Score 889; DB 13; Length 2860;  
Best Local Similarity 60.7%; Pred. No. 2,1e-211;  
Matches 1637; Conservative 0; Mismatches 960; Indels 102; Gaps 7;

QY 184 ACCAGTCCCTGGTCCCAACCGGACCTGCTTCCCACTTCTGTGGAGCCCGAGGATG 243  
Db 168 ACTCTCCCGTACGCGCAGCAGAGCCGCTGCCCTACTTCTGCGAGGAGCCACAGGAGC 227  
QY 244 TGTACATCGTCAAGAACAGCCAGTGTCTGTGTGTCAGAGCCGTGCGCCGACGCGAGA 303  
Db 228 CCTACATTTGTGAAGAACAGGCTGTGGAGCTCCGCTGCGCGCCTTCCCGCCACACAGA 287

QY 304 TCTTCTTCAAGTGCRAACGGGAGTGGTGGCCAGGTGGACCAAGTGCATCGAGCGGAGCA 363  
Db 288 TCTACTTCAAGTGCRAACGGGAGTGGTGGCCAGGTGGACCAAGTGCATCGAGCGGAGCC 347  
QY 364 CAGACGGGAGCAGTGGTGGAGCCGACCATGAGGTCCGCAATTAATGTCTCAAGCAGCAGG 423  
Db 348 TGGATGAGGCCACCGGTCTGCGGGTGGCGAGGTGCAGATCGAGGTGTCCGGCAGCAGG 407  
QY 424 TCGAAGAGTGTTCGGGCTGGAGGAATACTGTGCCAGTCCGTGGCATGGAGTCTCTCGG 483  
Db 408 TGGAGGAGCTCTTTGGGCTGGAGGATTAAGTGGTCCAGTGGCTGGAGTCTCCGCG 467  
QY 484 GCACCAACCAAGAGTGCAGAGGCTTACATCGCATACCAAGATTCGCAAGAACTTCGAGC 543  
Db 468 GCACCAACCAAGAGTGCAGAGGCTTACATCGCATACCAAGATTCGCAAGAACTTCGATC 527  
QY 544 AGGAGCGGTGGCCAAAGAGGTGTCCTGGAGCAGGGCATGCTGTGCTGCCCTGCCGTCAC 603  
Db 528 AGGAGCTCTGGGCAAGAGGTGCCCTCGGACCATAGAGTTCTCTGAGTGGCGCCGCG 587  
QY 604 CGGAGGGCATCCCTCCAGCGAGGTGGAGTGGCTCGGAACGAGGACCTGGTGGACCCGT 663  
Db 588 CGGAGGGGTGCTGTGGCGGAGTGGATGGCTCAAGAAATGAGGATGTCATCGACCCCA 647  
QY 664 CCTGGACCCCAATGTATATACATCAGCGGGAGCAGCCTGGTGGTGCACAGGCCCGCG 723  
Db 648 CCAGGACACCAACTTCTGTGCTCAGCATCGACCAACCTCATCTCCCGAGGCCCGCG 707  
QY 724 TTGCTGACACGGCCCAACTACCTCGGTGGCCCAAGAAATCTGTGCAGCTGCCCGCAGG 783  
Db 708 TGTGCGACACTGCCAACTATACCTGGTGGCCCAAGAAATCTGTGCACCAACGCCGAGCA 767  
QY 784 CTTCCGCTGTGTCTGTGACGAAAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 843  
Db 768 CCACTGCCACCGTCTCATCTGTACGTGAATGGCGGTGGTCCAGTGGGCGAGAGTGTCTAC 827  
QY 844 TCTGAGCGCCAGCTGTGGCGCGGTGGCAGAAAGGAGCCGAGCTGCACCAACCCCG 903  
Db 828 CTTGCTCCAAACCGCTGTGGCGAGGTGGCAGAAAGGAGCCGAGCTGCACCAACCCCG 887  
QY 904 CGCTCTCAACGGGGCGCTTCTGTGAGGGGCAGAAATGTCTCATGACGCCACCGTCTCT 963  
Db 888 CTCCACTCAACGGAGGGGCTTCTCGAGGGCCAGGCAATCCAGAA---GACCGCTGCA 944  
QY 964 TCTGCTGTGTCTGTGACCGCAGCTGGAGCCCGTGGAGCAAGTGGTGGGCTGTGGGC 1023  
Db 945 CCACCATCTGCCCAGTTCGATGGGGGTGGACGAGTGGAGCAAGTGGTGGTGGTGGTGG 1004  
QY 1024 TGGACTGCACCCACTGGCGGAGCGGTGAGTGTGTGACCCAGCACCCCGCAACGAGGGG 1083  
Db 1005 CTGAGTGTGCCCACTGGCGGTAGCGCGAGTGCATGGCGCCCGCCCGCCAGAACCGGAG 1064  
QY 1084 AGGAGTGCAGGGCAGCTGACCTGGACACCCCACTGTACCACTGACCTCTGTGTACACA 1143  
Db 1065 GTGACTGCAGCGGAGCGTGTCTGCACTTAAGAACTGCACAGATGGGTGTGCAATGCAAC 1124  
QY 1144 GTGCTTTCGGCCCTGAGGACGTGGCCCTCTATGTGGGCTCATGCGCGTGGCGGCTGTGC 1203  
Db 1125 TGGAGGCTTCAGGGGATGCGGCGCTGTATGCGGGCTGTGTGGTGGCCATCTTCGTGTG 1184  
QY 1204 TGGTCTGTGCTGTCTGTCTCTCATCTCTGTTATTGCGGAGAGGAGGGGTGGACT 1263  
Db 1185 TGGCAATCTCATGGCGTGGGGGTGGTGTGTATCCCGCCGCAACTGCGGTGACTTTCGACA 1244  
QY 1264 CAGATGTGGTGTGCTCTGTCATT---CTCACTCAGGCTTCCAGCCGTTCAGATCAAGC 1320  
Db 1245 CAGACATCACTGATCATCTGCTGCCCTGACTGGTGGTGTTCACCCCGTCAACTTTAAGA 1304  
QY 1321 CCAGCAAGCAGACAAACCCCACTGTCT-----CACCATCCAGCCGACCTTCAGACCA 1374  
Db 1305 CGGCAAGGCCCAGTAAACCCGAGCTCCTACACCCCTCTGTGCTCTCTGACCTGACAGCA 1364  
QY 1375 CCACCCTTACCAGGGCAGTCTGTCTCCCGGAGGATGGGCCCGACCCCAAGTTCACG 1434

Db 1365 GCGCGGCAATCTACCGCGGACCCGCTGATGCCCTGCAGACTCCACCGCAAAATCCCA 1424  
Qy 1435 TCACCAATGGGCACTGCTCAGCCCC----- 1460  
Db 1425 TGACCAACTCTCTCTGCTGGACCCCTTACCCAGCCCTTAAGGTCAAGGTCTACAGCTCCA 1484  
Qy 1461 -----CTGGGTGGCGGCGCC 1476  
Db 1485 GCACACCGGGCTCTGGCGCAGGCCCTGGCAGATGGGGCTGACCTGTGGGGGTCTTGGCGC 1544  
Qy 1477 ACACACTGCACACAGCTCTCCACTCTGAGGCGAGGAGTTGCTCTCCGCTCTCCA 1536  
Db 1545 CTGGGCAATACCTAGCAATTTGCGCGGGAACCCACTTCTGTGACCTGCGCAGGGCCA 1604  
Qy 1537 CCCGAACCTAC-----TTCCGCTCCCTGCCCCGAGGCAACCAATGACCTATGGGA 1590  
Db 1605 GCTCGGTTCCAGCAGCTCTTGGGCTGCGCCGAGACCCAGGAGCAGCGTCAGCGCA 1664  
Qy 1591 CTTTCAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCATCC 1650  
Db 1665 CTTTGGCTGCTGGGTGGAGGCTCAGCATCCCCGGCACAGGGGTGAGCTTGTGGTGC 1724  
Qy 1651 CCCAGATGCCATACCCGAGGAAGATCTATGAGATCTACCTCAGCTGCACAAAGCCGG 1710  
Db 1725 CCAATGGAGCCATCCCCAGGCAAGTTCTACGATGTATCTACTCATCAACAGGCAG 1784  
Qy 1711 AAGAGTGAAGTGGCCCTAGCTGTGTGCAGACCTGTGATGCCATCGTTAGCTGTG 1770  
Db 1785 AAGTACCTCGCGCTTTTCAAGAGGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844  
Qy 1771 GACCCCTGGCTGCTGTCTACCGGCGAGTCACTCTGCTATGACCACTGTGGGAGC 1830  
Db 1845 GACCACAGGCTCTGTGTGCGCGCCGCTATCCTCACCATGCCCCACTGTGCCGAAG 1904  
Qy 1831 CAGCCCTGACAGCTGGAGCCTCGGCTCCTCAAAAGCAGTCGTGCGAGGCGAGCTGGAGC 1890  
Db 1905 TCAGTCCCGTGAAGTGTCTTTCAGCTCAGACCCAGGCGCACAGGGCCACTGGAG- 1963  
Qy 1891 AGGATGTGTCACTTGGCGGAGGCGCCCTCCCACTTACTACTGCTGCTGAGG 1950  
Db 1964 --GAGTGTGACCTCGGATGAGGAGACCTGAACACACCTGTACTGCTGAGC 2021  
Qy 1951 CAGTGCCTGCTACGTCTTACCGAGCAGTGGCGGCTTTCCTGTGGGAGAGGCC 2010  
Db 2022 CCAGGCGCTGTCACTCTGCTGTGACGAGTGGCGACCTACGTTTACGCGGAGTCT 2081  
Qy 2011 TCAGCTGTGCTGCGGCAAGCGCTCAAGCTGCTTCTGTTTGGCGGCTGCGCTGACCT 2070  
Db 2082 ATTCCGCTCAGGAGTCAAGCGGCTCAGCTGGCGCTTTCGCGCCCGGCTCTGACCT 2141  
Qy 2071 CCTCGAGTACAACTCCGGGTCTACTGCTGATGACACCCACAGATGACTCAAGAGG 2130  
Db 2142 CCTGGAGTACAGCTCCGGGTCTACTGCTGGAGGACAGCGCTGTAGCACTGAAGGAGG 2201  
Qy 2131 TGGTGCAGTGGAGAGCAGCTGGGGGACAGCTGATCAGGAGCCACGGGTCTGCACT 2190  
Db 2202 TGTGAGTGGAGCGGATCTCTGGCGGATCTTGGTGGAGGAGCGCAACCGCTAATGT 2261  
Qy 2191 TCAAGGACAGTTACCAACCTCGGCTATPCCATCCACGATGTCGCGCTGCTGGA 2250  
Db 2262 TCNAGGACAGTTACCAACCTCGGCTCTCCCTCCATGACCTCCCGATGCCATTTGA 2321  
Qy 2251 AGAGTAACTCTTGTAGCTACAGGAGATCCCTTTTATATCATCTGGAATGGCAGC 2310  
Db 2322 GGAGCAAGTGTCTGGCGCAATATACAGGAGATCCCTTCTATCATATTTGAGTGGAGCC 2381  
Qy 2311 AGCGTACTTGCAGTCACTTCACTTCACTGAGGCTGTACGCCCCAGCACTAGTGA 2370  
Db 2382 AGAAGGCCCTCCACTGCACTTTTCACTTGGAGGAGCAGCTTGGCTCCACAGAGCTCA 2441  
Qy 2371 CCTGCAAGCTGTGGGTGTGSCAGGTGGAGGCGAGCGGCTTTCAGCATCACTTCA 2430

Db 2442 COTGCAAGATCTGCTGGCGCAAGTGAAGGGAGGCGCAGATATTCAGCTGCATCCA 2501  
Qy 2431 ACATCACAAG---GACAAGAAGTTTGTGAGCTGCTGGCTCTGAGAGTGAAGCGGGG 2487  
Db 2502 CTCTGGCAGACACACTGCTGGCTCCCTGGACACTCTCTGTCTCTGCCCTGGCAGACTG 2561  
Qy 2488 TCCAGCCCTGTGTGGGCCCCAGTGGCTTCAAGATCCCTTCTCATTCGGCAGAAATAA 2547  
Db 2562 TCACACCCAGCTGGGACTTATGCTTCAAGATCCCACTGTCCATCCGCCAGAGATAT 2621  
Qy 2548 TTTCCAGCTGGACCCACCCCTGTAGCGGGGTGCCAGCTGGGGGACTCTGGCCCAAGAAC 2607  
Db 2622 GCACAGCCTAGATCCCCCACTCACGGGGCAATGACTGGCGGATGTTAGCAGAGAAGC 2681  
Qy 2608 TCCACTCGACAGCCATCTCAGCTTCTTTCCTCAAGCCAGCCAGCCCAAGCATGATCC 2667  
Db 2682 TCTCTATGACCGGTACCTGAAATTAATTGCCCACCAAGAGGAGCCCGGTGTGATCC 2741  
Qy 2668 TCAACCTGTGGAGGCGGCGCACTTCCCCAAGCGGCAACCTCAGCCAGCTGGCTGAGCAG 2727  
Db 2742 TGGACTCTGGGAAGCTCTGCAGCAGGAGGAGTGGGACCTCAACAGCTTGGCGAGTGCCT 2801  
Qy 2728 TGGCTGGAAGTGGGCGCAGCCAGACGCTGGCTTCTTCAAGTGTGGAGGCTGAGTGTGA 2786  
Db 2802 TGGAGGAGATGGGCAAGAGTGAATGCTGTGGCTGTGGCCACCGAGCGGGGACTGCTGA 2860

## RESULT 13

US-10-218-779-1

; Sequence 1, Application US/10218779  
; Publication No. US2004002922A1

## GENERAL INFORMATION:

; APPLICANT: Edinger, Shlomit  
; APPLICANT: Machougall, John  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Grosse, William  
; APPLICANT: Alsbrook II, John  
; APPLICANT: Lepley, Denise  
; APPLICANT: Rieger, Daniel  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Casman, Stacie  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Tchernev, Vellizar  
; APPLICANT: Vernet, Corine  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Miller, Charles  
; APPLICANT: Gangoli, Bsha  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-214  
; CURRENT APPLICATION NUMBER: US/10/218,779  
; CURRENT FILING DATE: 2002-08-14  
; PRIOR APPLICATION NUMBER: 60/253,834  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: 60/250,-926  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/264,180  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/313,656  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: 60/327,456  
; PRIOR FILING DATE: 2001-10-05

```

; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2860
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-218-779-1

```

Query Match	30.9%;	Score 889;	DB 13;	Length 2860;
Best Local Similarity	60.7%;	Pred. No. 2.1e-211;		
Matches 1637;	Conservative	0;	Mismatches 960;	Indels 102;
	Gaps	7;		
QY	184	ACCCAGTGCCTGGTGCCAAACCGGACCTGCTTCCCACTTCTCTGGTGAAGCCCGAGGATG	243	
DB	158	ACTCCTTCGGTCAAGCGCCAGAGAGCCGCTGCCTACTTCTCGAGGAGCACAGGACG	227	
QY	244	TGTACATCGTCAAGAAACAGCCAGTGTCTGTGTGCAAGGCCGTGCGCCCGCCACGAGCA	303	
DB	228	CCTACATTGTGAAGAAACAGCCCTGTGGAGCTCCGCTGCGCGCCCTTCCCGCCACACAGA	287	
QY	304	TCCTTTCAAGTGCMAACGGGAGTGGGTGGCCGAGGTGGACCAACGTGATCGAGCGCAGCA	363	
DB	288	TCTACTTCAAGTGCMAACGGCGAGTGGGTGACGCCAAGCACCGTACACAGGAAGGCC	347	
QY	364	CAGACGGAGCAGTGGTGAGCCGACCATGGAGGTCGGCAITTAATGTTCTCAAGGCAGCAGG	423	
DB	348	TGGATGAGGCCACCGGTTCTGGGTGCGGAGGTGCAGATCGAGGTGTCCGCGCAGCAGG	407	
QY	424	TCGAGAAGGTGTTCCGGCTGGAGGAATACTGGTGCCAGTGGCTGGCATGGAGCTTCCTCGG	483	
DB	408	TGGAGGAGCTCTTTGGGCTGGAGGATTACTGTTGCCAGTGGTGGCTTGGAGCTCCGCGG	467	
QY	484	GCACCAACAGAGTCAGAAAGCCCTACATCCGCATAGCCAGATTGCGCAAGAACTTCGAGC	543	
DB	468	GCAACCAAGAGTCGCGAGCCCTACGTCGCGATCGCCTACCTCGCAAGAACTTCGATC	527	
QY	544	AGGAGCGCTGGCCAAAGAGAGTGCCCTGGAGCAGGCATCGTCTGCCCTGCGGTCAC	603	
DB	528	AGGAGCCTCTGGGCAAGGAGTGCCCTTGGACATGAGGTTCTCTCGATGCGCGCCCGC	587	
QY	604	CGGAGGGATCCCTCCAGCCAGGTGAGTGGTTCGGAAACGAGACCTTGGTGAGACCCGT	663	
DB	588	CGGAGGGGTGCTTGTGGCGAGGTGGAATGGCTCAAGATGAGGATGTATCGACCCCA	647	
QY	664	CCCTGGACCCCAATGTATACATCAGCGGAGGACACGCTGGTGGTTCGAAGAGCCCGCC	723	
DB	648	CCGAGGACCAACTTCTGCTACCATCGACCAACACCTCATCATCCGACGAGCCCGCC	707	
QY	724	TTGCTGACCGGCCAACTACACCTTCGCTGGCCAGAAACATCGTGGACGTCGCGCGAGCG	783	
DB	708	TGTCGGACATCGCCAACTATACCTGCGTGGCCCAAGAACATCGTGGCCAAACGCGGAGCA	767	
QY	784	CCTTCGGCTGTGTCATCGTCTAGTGAAACGCTGGGTGGTTCGACGTGACCGAGTGTTCG	843	
DB	768	CCAATGCCACCGCTATCGTCTACGTGAATGGCGGCTGGTCCAGCTGGCAGAGTGTGTTC	827	
QY	844	TCTGCAAGCGCCAGCTGTGGCGGGCTGGCAGAAACGGAGCCGAGCTGCACCAACCCCG	903	
DB	828	CCTGTCTCAACCGGTGTGGCCGAGGCTGGCAGAAAGCGCACCCGGACCTGCACCAACCCCG	887	
QY	904	CGCCTCTCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCATGACCGCACCGTCTCCT	963	
DB	888	CTCCACTCAACGGAGGGGCTTCTGTGGAGGGCCAGGCATTCAGAA----GACCGCTGCA	944	
QY	964	CTCTGCTTGTCTCTGTGGACGGCAGCTGGAGCCCGTGGAGCAAGTGGTGGGCTGTGGGC	1023	
DB	945	CCACCATCTCCCCAGTGCATGGGCGGTGGACGAGTGGAGCAAGTGTGAGCTGTCAGCA	1004	
QY	1024	TGGACTGCACCCACTGGCGGAGCCGTGAGTGTCTTGACCCAGACCCCGCAACGGAGGG	1083	
DB	1005	CTGAGTGTGCCACTGGCGTGTAGCCGAGTGTGCATGGCGCCCCCCCCACCCAGACGGAGGCC	1064	
QY	1084	AGGAGTCCGAGGCACTGACCTGGACACCGGCAACCTGTACAGTGAACCTCTGTGTACACA	1143	

Db 2142 CCCTGGAGTACAGCCTCCGGGTCTACTGCTGGAGGACAGGCTCTAGCACTGAAGGAGG 2201  
QY 2131 TGGTCCAGCTGGAGAGCAGCTGGGGGACAGCTGATCCAGAGGACAGGCTCTGCACT 2190  
Db 2202 TGCTGGAGCTGGAGCGGACTCTGGGCGGATACTTGGTGAGGAGCGGAAACCGCTAAATGT 2261  
QY 2191 TCAGGACAGTTACACACACCTGCGCCCTATCCATCCAGATGTGCCAGCTCCCTGTGA 2250  
Db 2262 TCAGGACAGTTACACACACCTGCGCCCTCTCCCTCCATGACTCCCTCCCATGCCCATTTGA 2321  
QY 2251 AGAGTAAGCTCCCTGCTACAGTACACAGGAGATCCCTTTTATCACATCTGGAATGACGC 2310  
Db 2322 GGAGCAAGCTGTGCCAAATACACAGGAGATCCCTTCTATCACATTTGGAGTGGCAGCC 2381  
QY 2311 AGCGTACTTGACCTGACCTTACCTGGAGCGTGTAGCCCCAGCACTAGTACCTGG 2370  
Db 2382 AGAAGCCCTCCACTGCACTTTCACCTGGAGAGGACAGCTTGGCTCCACAGAGCTCA 2441  
QY 2371 CCTGCAAGCTGGGTGGGAGGTGGAGGGGAGCGGAGAGCTTTCAGCATCAACTTCA 2430  
Db 2442 CCGCAAGATCTGGTGGGCAAGTGGAGGGGAGGCGGAGATATTCAGCTGCATACCA 2501  
QY 2431 ACATCACCAAG---GACACAAGTTTGTGAGCTGCTGCTCTGGAGAGTGAAGCGGGG 2487  
Db 2502 CTCGCGCAGAGACACTGCTGGCTCCCTGGACACTCTCTGCTCTGCCCTGGCAGCACTG 2561  
QY 2488 TCCAGCCTGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCTATTGGCGAGAGATRA 2547  
Db 2562 TCACCACCAGCTGGGACCTTATGCTTCAAGATCCCACTGCTCCATCCGCGCAGAGATAT 2621  
QY 2548 TTTCCAGCTGGACCCACCTCTAGGCGGGGTGCCACTGGCGGACTCTGGCCCGAGAAC 2607  
Db 2622 GCAACAGCTAGTGCCTCCCACTACGCGGGAATGACTGGCGGATTTAGCACAGAGC 2681  
QY 2608 TCACCTGGACAGCCATCTAGCTTTTGTCTTCCAAAGCCCGAGCCCAAGCCCATATCC 2667  
Db 2682 TCTCTATGACCGGTACCTGTAATTACTTTGCCAACAAAGCGAGCGCGGTGTATCC 2741  
QY 2668 TCACCTGGAGGCGGCGGACTTCCCAAGCGGCAACTCAGCCAGCTGGCTGCGAGCAG 2727  
Db 2742 TGGACCTCTGGGAAGCTCTGCGCAGAGCAGCATGGGAGACCTCAACAGCCTGGGAGTGCCT 2801  
QY 2728 TGGCTGACTGGCCAGCAGCAGCTGGCTCTTTCACAGTGTGGAGGTGAGTGTCTGA 2786  
Db 2802 TGGAGGAGTGGCGAAGAGTGAATGCTGCTGGTGTGCCACCAGCGGGAGCTGCTGA 2860

## RESULT 14

US-10-087-684-3  
; Sequence 3, Application US/10087684  
; Publication No. US20040029116A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David J.  
; APPLICANT: Grosse, William M.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Casman, Stacie, J.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Boldog, Ferenc L.  
; APPLICANT: Li, Li  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Mishra, Vishnu  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Malyankar, Uriel M.

; APPLICANT: Guo, Xiaojia  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Gangoli, Esha A.  
; TITLE OF INVENTION: PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-214 CIP  
; CURRENT APPLICATION NUMBER: US/10/087,684  
; CURRENT FILING DATE: 2003-03-10  
; PRIOR APPLICATION NUMBER: 60/253,834  
; PRIOR FILING DATE: 2000-11-29  
; PRIOR APPLICATION NUMBER: 60/250,926  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: 60/264,180  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/313,656  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: 60/327,456  
; PRIOR FILING DATE: 2001-10-05  
; NUMBER OF SEQ ID NOS: 220  
; SOFTWARE: CuraseqList version 0.1  
; SEQ ID NO 3  
; LENGTH: 2860  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (59)..(2857)  
US-10-087-684-3

Query Match 30.8%; Score 887.4; DB 13; Length 2860;

Best Local Similarity 60.6%; Pred. No. 5.4e-211;

Matches 1636; Conservative 0; Mismatches 961; Indels 102; Gaps 7;

QY 184 ACCAGTGCCTGGTGCACACCGGACCTGCTTCCCACTCTCTGGTGGAGCCGAGGATG 243  
Db 168 ACTCCTTCCCGTCAGCGCCAGCAGCGCGCTCTCTCTGAGGAGCCACAGGAGC 227  
QY 244 TGTACATCTCAAGAACAGCCAGTGTCTTGTGTGCAAGCGCGTCCCGCCACGCAGA 303  
Db 228 CCTACATTTGAAGAACAGCTTGGAGCTTTCGCTGCGCGCGCTTCCCGCCACACAGA 287  
QY 304 TCTTCTTCAAGTGCACACGGGAGTGGGTGCGCAGGTGACCAACGATGAGCGCGAGCA 363  
Db 288 TCTACTTCAAGTGCACACGGGAGTGGGTGAGCGAGAACGACACGTCACACAGGAAGGCC 347  
QY 364 CAGACGGGAGCAGTGGTGAAGCCAGCATGAGGTCCGANTTATGTCTCAAGGSCAGCAGG 423  
Db 348 TGGATGAGGCCACCGGCTGCGGGTGCAGATGAGGTGTGCGCGAGCAGG 407  
QY 424 TCGAGAAGGTGTTCCGGCTGGAGGATACTGGTGCAGTGCCTGGCATGGAGCTCTCTCGG 483  
Db 408 TGGAGGAGCTCTTTGGGCTGGAGGATTAATGGTGCAGTGCCTGGAGCTCCGAGC 467  
QY 484 GCACCAACAGAGTCAGAAAGGCTACATCCGATAGCCAGATTGCGGAAGAACTTCGAGC 543  
Db 468 GCACCAACAGAGTGCAGGCTTACGTCCGATCCCTACCTGCGCAAGAACTTCGATC 527  
QY 544 AGGAGCGCTGGCCAGAGGAGTGTCCCTGAGCAGGAGCATCGTGTGCGCTGCGCTCCAC 603  
Db 528 AGGAGCCTTGGGCAAGGAGGTGCGCTGAGCATGAGGTTCCTGCGAGTGGCGCGCC 587  
QY 604 CGGAGGGCATCCCTCCAGCGAGGTGGGTGCGTCCGGAACGAGGACCTTGGTGGACCCGT 663  
Db 588 CGGAGGGGTGCTGTGGCGGAGGTGGAATGGCTCAAGATGAGGATGTCATGACCCCA 647  
QY 664 CCTGGAACCCCAATGATATCATACCGCGGAGCAGACGCTGGTGGTGGAGCAGCCCGCC 723  
Db 648 CCCAGGACACCAACTTCTCTGCTCACCATCGACACCAACCTCATCATCCGAGCGCCGCC 707  
QY 724 TTGCTCAGACGGCCAACTACCTCTGCTGSCCAAGAACATCTGTGGCAGCTGCGCGCAGCG 783  
Db 708 TGTCCGACACTGCCAACTATACCTGCTGTGCCAAGAACATCTGTGGCCAAACGCGCGGAGCA 767





GENERAL INFORMATION:  
APPLICANT: Edinger, Shlomit  
APPLICANT: MacDougall, John  
APPLICANT: Millet, Isabelle  
APPLICANT: Ellerman, Karen  
APPLICANT: Stone, David  
APPLICANT: Gerlach, Valerie  
APPLICANT: Grosse, William  
APPLICANT: Alsobrook II, John  
APPLICANT: Lepley, Denise  
APPLICANT: Rieger, Daniel  
APPLICANT: Burgess, Catherine  
APPLICANT: Casman, Stacie  
APPLICANT: Spytek, Kimberly  
APPLICANT: Boldog, Ferenc  
APPLICANT: Li, Li  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Mishra, Vishnu  
APPLICANT: Patturajan, Meera  
APPLICANT: Shenoy, Suresh  
APPLICANT: Rastelli, Luca  
APPLICANT: Tchernev, Velizar  
APPLICANT: Vernet, Corine  
APPLICANT: Zerhusen, Bryan  
APPLICANT: Malyankar, Uriel  
APPLICANT: Guo, Xiaojia  
APPLICANT: Miller, Charles  
APPLICANT: Gangolli, Bsha  
TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
FILE REFERENCE: 21402-214  
CURRENT APPLICATION NUMBER: US/10/218,779  
PRIOR FILING DATE: 2002-08-14  
PRIOR APPLICATION NUMBER: 60/253,834  
PRIOR FILING DATE: 2000-11-29  
PRIOR APPLICATION NUMBER: 60/250,-926  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: 60/264,180  
PRIOR FILING DATE: 2001-01-25  
PRIOR APPLICATION NUMBER: 60/313,656  
PRIOR FILING DATE: 2001-08-20  
PRIOR APPLICATION NUMBER: 60/327,456  
PRIOR FILING DATE: 2001-10-05  
NUMBER OF SEQ ID NOS: 216  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 3  
LENGTH: 2860  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-218-779-3

Query Match 30.8%; Score 887.4; DB 13; Length 2860;  
Best Local Similarity 60.6%; Pred. No. 5.4e-211;  
Matches 1636; Conservative 0; Mismatches 961; Indels 102; Gaps 7;

QY 184 ACCAGTGTCTGGTCCAAACCCGAGCTGCTTCCCACTTCCGTGGAGCCGAGATG 243  
Db 168 ACTCTTCCCGTCAGCGCCAGCAGACCGCTGCGCTTCTCTCGAGAGCCACAGAG 227

QY 244 TGATACGTCTCAAGAAACAGCCAGTGTCTTGTGTGCAAGGCGGTGCCCGCCACGAGA 303  
Db 228 CCTACATTGTGAAGAACAGCTGTGGAGCTTCGCTGCGCGCTTCCCGCCACACAGA 287

QY 304 TCTTCTTAAGTCCAAAGGGAGTGGTGGCGCAGGTGGACCGAGGTGATCGAGCGAGCA 363  
Db 288 TCTACTTCAAGTCAACGGGGAGTGGGTGAGCCAGAACAGCACGTCACACAGGAAGGCC 347

QY 364 CAGACGGGACAGTGTGAGCCAGCCACCATGAGGTGCCCATTAATGTCTCAAGGCGAGG 423  
Db 348 TGGATGAGCCACCGGCTCGGGTGGCGAGGTGCAGATCGAGGTGTCGGCGAGCAGG 407

QY 424 TCGAGAAGTGTTCGGGCTGGAGGAATACCTGGTSCCAGTGCCTGGCATGAGTCTCTCG 483  
Db 408 TGGAGGAGCTCTTGGGCTGGAGGATTACTGGTSCCAGTGCCTGGGCTGGAGTCCGCG 467

QY 484 GCACCAACAGAGTCAGAGGCGCTATCCGATAGCCAGATTGGCAAGAACTTCGAGC 543  
Db 468 GCACCAACAGAGTCGCGAGCGCTACGTCGCGATCGCTACTCTGCGCAAGAACTTCGATC 527

QY 544 AGGAGCGCTGGCCAAAGGAGGTGTCCCTGGAGCAGGCGATCGTGTGCCCTGCCCTCCAC 603  
Db 528 AGGAGCTCTGGGCAAGGAGGTGCCCCCTGGACCATGAGTTCTCTCGAGTCCGCGCCGC 587

QY 604 CGGAGGCGATCCCTCCAGCCGAGGTGGTCTCGGAAACAGAGACCTGTGTGGACCCGT 563  
Db 588 CGGAGGCGGTGCTGTGGCCGAGGTGGAAATGGCTCAAGAAATGAGGATGTTCATCGACCCCA 647

QY 664 CCCTGGACCCCAATGTATACATCAGCGGAGCAGAGCTGTGTGGCAGAGGCCGCGC 723  
Db 648 CCCAGGACACCAACTTCTGTCTCACCATCGACCAACCTCATCTCCGCAAGGCCGCGC 707

QY 724 TTGCTGACACGCGCAACTACACTGCTGGCGCAAGAAACATCGTGGCACTGCCGCGAGCG 783  
Db 708 TGTGGACACTGCCAACTATACCTGGTGGCCCAAGAAACATCGTGGCAAAACGCGGAGCA 767

QY 784 CTTCCGCTGTCTCATCTGTACGTGAACGTTGGTGGTGTGTCGACGTGGACCGAGTGGTCCG 843  
Db 768 CCACCTGCCACCGCTCATCTGTACGTGAATGGCGGCTGTGTCCAGCTGGGCGAGATGGTTCAC 827

QY 844 TCTGAGCGCCAGCTGTGGCGCGCTGSCAGAAACGAGCGGAGCTGCACCAACCCCG 903  
Db 828 CTTGCTTCCAAACCGCTGTGGCCGAGCTGSCAGAGGCGACCCGCGACCTGTCACCAACCCCG 887

QY 904 CGCTCTCAACCGGGGCGCTTCTGTGAGGGGCGAAGTTCATGACCGCACCGTCTCTCT 963  
Db 888 CTCACACTCAACGAGGGGCTTCTCGAGGGCGCAGGCATTCAGAA---GACCGCTGCA 944

QY 964 CTCTGCTGTCTGTGAGCGGAGCTGGAGCCCGTGGAGCAAGTGTGGCTGTGGGCG 1023  
Db 945 CCACCATCTGCCCGAGTCGATGGGCGTGGACGAGTGGAGCAAGTGTGAGCTCGAGCA 1004

QY 1024 TGGAGTGCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGG 1083  
Db 1005 CTGAGTGTGCCCACTGGCGGTAGCCGAGTGTATGGCGCCCGCCACCCAGAGCGAGGCC 1064

QY 1084 AGGAGTGCACGAGGCACTGACCTGGACACCCCGCAACTGTACCAAGTACCTCTGTGTACACA 1143  
Db 1065 GTGACTGCAGCGGGAGCGTGTCTGACTCTAAGAACTGCACAGATGGGCTGTGCATGCAAC 1124

QY 1144 GTGCTTGTGGCCCTGAGGAGTGGCCCTCTATGTGGGCTCATCGCGGTGGCGTCTGCC 1203  
Db 1125 TGGAGGCGCTCAGGGGATGCGGCGCTGTATGCGGGGCTCGTGTGGCCATCTTCGTGGTGG 1184

QY 1204 TGGTCTGTCTGTCTGTCTCTCATCTCTGTTTATTCGCGGAGAGAGGGGCTGGACT 1263  
Db 1185 TGGCAATCTCATGGCGGTGGGGGTGGTGTGTACCGCGCAACTCCCGTGTGACTTCGACA 1244

QY 1264 CAGATGTGGCTGATCTGTCCATTT---CTCACTCAGGTTCAGCGCGCTGAGCATCAAGC 1320  
Db 1245 CAGACATCACTGACTCATCTGTGCGCTGACTGTGGTGTTCACACCCCGTCAACTTTAAGA 1304

QY 1321 CCAGCAAGAGACACACCCCGCATCTGCT---CACCATCCAGCGGAGCTCAGACCA 1374  
Db 1305 CGGCAAGGGCCAGTAACCCGCGAGCTCTACACCCCTCTGTGCTCTCTGACTCTGACGACCA 1364

QY 1375 CCACCACTTACCAAGGCGAGTCTGTCTGTCGCGGAGGATGGGCGGAGCCCAAGTTCAGC 1434  
Db 1365 GCGCGGCGCATCTACCGGGACCGGTGTATGCCCTGCGAGGACTCCACCGCAAAATCCCA 1424

QY 1435 TCAACCAATGGGCACTGTCTCAGCCCG---CTGAGGCTTAAAGGTCAAGGTCTACAGTCTCA 1460  
Db 1425 TGACCAACTCTCTCTGTGGACCCCTTACCAGCCCTTAAAGGTCAAGGTCTACAGTCTCA 1484

QY 1461 -----CTGGGTGGCGCGCGC 1476  
Db 1485 GCACCAAGGCTCTGGGCGAGGCGCTGGCAGATGGGGCTGACCTGCTGGGGGTCTTGGCGC 1544

QY 1477 ACACACTGACACAGCTCTCCACACTCTGAGCGGAGAGTTCGTCTCCCGCTCTCCA 1536  
Db 1545 CTGGCACATACCTTAGCAATTCGCGCGGACACACCTCTCTGCACTCTCGGAGGCCA 1604  
QY 1537 CCAGAACTAC-----TTCCGCTCCCTGCGCGGAGCACACAGCAATGACCTATGGGA 1590  
Db 1605 GCCTCGGTTCGCCAGAGCTCTTGGGCTGCGCGGAGACCCAGGGAGCAGGGTCAGCGGCA 1664  
QY 1591 CTTCAACTCTCTCGGGGCGGCTGATGATCCCTAATACAGTATACAGCTCTCTCATCC 1650  
Db 1665 CTTTGGCTGCTGGTGGAGGCTCAGCATCCCGGCAAGGGGTGAGCTTGGTGC 1724  
QY 1651 CCCGATGCTATACCCCGAGGGAAGATCTATGAGATCTACTCTCAAGCTGCAAGCGG 1710  
Db 1725 CCAATGGAGCAATTCGCCAGGCAAGTCTACGAGATGATCTACTCATCAACAGGCAG 1784  
QY 1711 AAGACCTGAGTTGCCCTAGCTGGCTGTCAGACCTCTGCTGAGTCCCATCGTTPAGTGTG 1770  
Db 1785 AAGTACCTGCGGCTTTCAAGAGGAGACCCAGACAGTATTGAGCCCTCGGTGACCTGTG 1844  
QY 1771 GACCCCTGCGCTCTGCTCACCGCGCCAGTCACTCTGCTATGAGCACTGTGGGAGC 1830  
Db 1845 GACCCACAGGCTCTCTGCTGTGCGCGCCGCTCATCTCAACATGCCCCACTGTGCCGAG 1904  
QY 1831 CCAGCCCTGACAGCTGAGGCTCGGCTCAAAAAGCAGTCTGTGCGAGGCGAGCTGGGAGC 1890  
Db 1905 TCAGTGCCCGTGACTGATCTTTACGCTCAAGACCCAGGCCACACAGGGCCACTGGGAG- 1963  
QY 1891 AGGATGTGCTGACCTGGCGAGGAGGCGCCCTCCCACTCTACTCTCCAGCTGGAGG 1950  
Db 1964 --GAGTGTGTACCTGGATGAGGAGACCTGAAACACACCTCTACTGCCAGCTGGAGC 2021  
QY 1951 CCAGTGCCTGCTACGCTTTACCGAGAGCTGGCGGCTTTGCCCTGGTGGGAGGCCCC 2010  
Db 2022 CCAGGSCCTGTCAATCTGCTGGACCAGCTGGGACACTACGTGTTACGGGCGAGTCT 2081  
QY 2011 TCAGGTGGCTGCGCCAAAGCGCTCAAGCTGCTTTCTGTTGCGCGGTGGCTGCACCT 2070  
Db 2082 ATTCCCGCTCAGCAGTCAAGCGGCTCCAGCTGGCGGCTCTTCCGCCCGCCCTCTGCACCT 2141  
QY 2071 CCCTCGAGTACAACTCCGGGTCTACTGCTGCTCATGACACCCACGATGCTCAAGGAGG 2130  
Db 2142 CCCTGGAGTACAGCTCCGGGTCTACTGCTGGAGGACACGCTGAGCACTGAAAGGAGG 2201  
QY 2131 TGTTCAGCTGGAGAGCAGCTGGGGGAGCAGCTCATCCAGAGCAGCGGTCTGCACT 2190  
Db 2202 TGCTGGAGCTGGAGCGGACTTGGCGGGATCTTGGTGGAGGAGCGGAAACCGCTTAATGT 2261  
QY 2191 TCAGGACAGTTACACAACTGCGCTTATCCATCCAGATGTGCCAGCTCCCTGTGGA 2250  
Db 2262 TCAAGGACAGTTACCAAACTGCGGCTCTCCCTCCATGACCTCCCCCATGCCCATTTGA 2321  
QY 2251 AGAGTACGCTTGTGCTACAGGAGATCCCTTTTATCAATCTGGAATGGCAGCC 2310  
Db 2322 GGAGCAGCTGTGGCCAAATACCAAGGAGATCCCTTCTATCAATTTGGAGTGGAGCC 2381  
QY 2311 AGCGGTACTTGCATGCACTTCACTTCACTGGAGGCTGTCAGCCCCAGCACTAGTACCTGG 2370  
Db 2382 AGAGGCCCTCCATGCACTTTCACCTTGGAGAGCAGCTTGGCCCTCCACAGAGCTCA 2441  
QY 2371 CCTGCAAGCTTGGGTGGGAGGTGGAGGGGAGCGAGGAGCTTTCAGCATCAACTTCA 2430  
Db 2442 CCTGCAAGATCTGCGTGGGCAAGTGGAGGAGGGGCGCAGATATTCAGCTGCATACCA 2501  
QY 2431 ACATCACCAAG---GACACAAGTTTGTGCTGAGCTGTGGCTCTGGAGAGTGAAGCGGGG 2487  
Db 2502 CTCCTGGAGAGACACCTGCTGGCTCCCTGGGACACTCTCTGCTCTGCCCTGGCAGCACTG 2561  
QY 2488 TCCAGAGCCTGGTGGGCCCCAGTGCCTTCAAGATCCCTTCTCATTCGGCAGAGATAA 2547  
Db 2562 TCACACCCAGCTGGGACCTTATGCTTCAAGATCCCACTGTCCATCGGCCAGAGATAT 2621  
QY 2548 TTTCCAGCTGGAACCCACCTGTAGGGGGGTGCGGACTGGCGGACTCTGGCCCCAGAAAC 2607

Db 2622 GCACAGCCTAGATGCCCCCAACTACGGGGCAATGACTGGCGGATGTTAGCACAGAAGC 2681  
QY 2608 TCACACTGGACAGCCATCTCAGCTTCTTTGCCCTCCAAAGCCAGCCGCCACAGCCATGATCC 2667  
Db 2682 TCTCTATGGACCGGTACTTGATTACTTTGCCACCAAGAGGAGCCCCACGGGTGTGATCC 2741  
QY 2668 TCACCTGTGGGAGGCGCGCACTTCCCCAAAGGCAACCTTCAGCCAGCTGGCTGCAGCAG 2727  
Db 2742 TGGACCTCTGGGAAGCTCTGCAGCAGGACGATGGGACCTCAACAGCCTGGCCGAGTGCCT 2801  
QY 2728 TGGCTGGAATGGGCAAGCCAGAGCTGGGCTCTTACAGTGTGGAGGCTGAGTGTGA 2786  
Db 2802 TGAGGAGATGGGCAAGAGTGAGATGCTGTGGCTGTGGCCACCGAGCGGGGACTGTCTGA 2860

Search completed: October 5, 2004, 21:09:21  
Job time : 1348 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 5, 2004, 15:09:57 ; Search time 7172 Seconds

(without alignments)

11995.670 Million cell updates/sec

Title: US-09-970-944-1

Perfect score: 2881

Sequence: 1 agctgggtcgggtgag.....ccttccacacccggggaga 2881

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:

1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estin.\*  
4: em\_estmu.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_htc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hrc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_mam.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	907.8	31.5	2802	29	AY406491	AY406491 Homo sapi
2	898.6	31.2	2791	29	AY406493	AY406493 Mus muscu
3	871.6	30.3	3790	11	AK031655	AK031655 Mus muscu
4	818.4	28.4	3866	11	AK018177	AK018177 Mus muscu

5	763.2	26.5	2532	29	AY411747	AY411747 Homo sapi
6	763.2	26.5	2802	29	AY406492	AY406492 Pan trogl
7	756.2	26.2	1034	12	BI758231	BI758231 603029876
8	726.4	25.2	2532	29	AY411749	AY411749 Mus muscu
9	726.2	25.1	909	13	BX364574	BX364574 BX364574
10	722.2	25.1	788	9	AI951556	AI951556 wv36f04.x
11	711.2	24.7	1532	11	BC033727	BC033727 Homo sapi
12	710	24.6	818	12	BI818609	BI818609 603033362
13	709.8	24.6	843	13	BX348193	BX348193 BX348193
14	700.2	24.3	939	13	BX345406	BX345406 BX345406
15	679.4	23.6	1040	13	BX345407	BX345407 BX345407
16	664.2	23.1	1201	13	BX422753	BX422753 BX422753
17	664	23.0	889	13	BO691915	BO691915 AGENCOURT
18	653.6	22.7	934	10	BF311804	BF311804 601897316
19	648.6	22.5	2532	29	AY411748	AY411748 Pan trogl
20	646.4	22.4	859	10	BF311896	BF311896 601897733
21	644.2	22.4	874	13	BO689148	BO689148 AGENCOURT
22	604.4	21.0	756	13	BU612387	BU612387 UI-M-EWO-
23	596.4	20.7	662	9	AL516580	AL516580 AL516580
24	579	20.1	2775	29	AY401471	AY401471 Mus muscu
25	573.4	19.9	604	10	BE314370	BE314370 601147261
26	573.4	19.9	2775	29	AY401469	AY401469 Homo sapi
27	554.6	19.3	1175	10	BF530640	BF530640 602071931
28	548	19.0	788	14	CA317532	CA317532 UI-M-FW0-
29	539.4	18.7	678	14	CA749784	CA749784 UI-M-FD0-
30	524.4	18.2	824	12	BI737024	BI737024 603360874
31	493	17.1	1030	12	EM479478	EM479478 AGENCOURT
32	471.2	16.4	540	12	BI467799	BI467799 389583 MA
33	467.8	16.2	917	13	BX453456	BX453456 BX453456
34	463.6	16.1	1028	13	BX331575	BX331575 BX331575
35	449.8	15.6	987	13	BX452510	BX452510 BX452510
36	445	15.4	2507	29	AY401470	AY401470 Pan trogl
37	433	15.0	471	13	BX282095	BX282095 BX282095
38	430.6	14.9	675	14	CA315487	CA315487 UI-M-FW0-
39	424.4	14.7	843	12	BG913440	BG913440 602811321
40	424	14.7	780	12	BG470852	BG470852 602511705
41	416	14.4	718	14	CA318180	CA318180 UI-M-FW0-
42	415.4	14.4	1147	13	BU840446	BU840446 AGENCOURT
43	413	14.3	960	12	BG329812	BG329812 602429458
44	411.6	14.3	1111	12	BG298307	BG298307 602397080
45	411.4	14.3	749	14	CF735417	CF735417 UI-M-HB0-

## ALIGNMENTS

RESULT 1	AY406491	2802 bp	DNA	linear	GSS 15-DEC-2003
LOCUS	AY406491	Homo sapiens	UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,		
DEFINITION	AY406491	Genomic survey sequence.			
ACCESSION	AY406491.1	GI:39762465			
VERSION	GSS.				
KEYWORDS	Homo sapiens (human)				
SOURCE	Homo sapiens				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
REFERENCE	1 (bases 1 to 2802)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Infering nonneutral evolution from human-chimp-mouse orthologous gene trios				
JOURNAL	Science 302 (5652), 1960-1963 (2003)				
PUBMED	14671302				
REFERENCE	2 (bases 1 to 2802)				
AUTHORS	Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.				
TITLE	Direct Submission				

JOURNAL	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA									
COMMENT	This sequence was made by sequencing genomic exons and ordering them based on alignment.									
FEATURES	Location/Qualifiers									
source	1..2802									
gene	/organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" <1..2802 /gene="UNC5C" /locus_tag="HCM2575"									
ORIGIN										
Query Match	31.5%; Score 907.8; DB 29; Length 2802;									
Best Local Similarity	60.3%; Pred. No. 3.9e-142;									
Matches 1660; Conservative	0; Mismatches 1014; Indels 78; Gaps 7									
Qy	98	GCCCGGCGCTGTGGCAGCGCTCTCTGGGCATAGTCTCTCGCCGCTTGGCTCCGCGGCTCGGG	157							
Db	66	GCTGTGCTACCTGCTCCCTGGCCCTGCTCAGCGCCAGCGCACTGGCTCGCGGCCAAGA	125							
Qy	158	TGCCCAGCAGAGTGCCACCGGTGGCCAAACCCAGTGCTGGTGCCAAACCCGAGACTGCTTCC	217							
Db	126	TGATGACTTTTTCATGAACCTCCCAAGAACTTTTCTCTCTGATCCACCTGAGCCTCTGCC	185							
Qy	218	CCACTTCTGTGTGAGCGCGGAGGATGTATCATGCTCAAGAACCAAGCCAGTGCTCTTGT	277							
Db	186	ACATTTCTTTATGTAGCTTGAAGAAGCTTATATTGTGAAGAAATAGCTGTGAACCTGTA	245							
Qy	278	GTGCAAGGCCGTGCCCGCCACGCGAGATCTTCTTCAAGTGCAAACGGGGAGTGGTGCGCA	337							
Db	246	CTGTAAAGCAAGCCCTGCCACCCAGATCTATTTTCAAGTGTAAATAGTGAATGGTTCATCA	305							
Qy	338	GGTGCACCATGTATCGAGCGCAGCAGCAGACGGGAGCAGTGGTGTAGCCGACCATGGAGT	397							
Db	306	GAAGGACACCATAGTAGATGAAGAGTAGATGAAACTTCCCGGTCTCATTTGTCGGGAAAT	365							
Qy	398	CCGCATTAAATGTCTCAAGGCACGAGTTCGAGAAAGTGTTCGGGCTTGGAGGAATATCTGGT	457							
Db	366	GAGCATTCAGATTTCGCCCCAGAGTGAAGAACTCTTTTGGACCTGAAGATTACTGGTG	425							
Qy	458	CCAGTCGTGGCATGGAGCTCCTCGGGCACACCAAGAGTCAAGAGCCATCATCCGCAT	517							
Db	426	CCAGTGTGTGCGCTGGAGCTCCGCGGGTACCAACAAGAGCGGAAAGGCGTATGTGCGCAT	485							
Qy	518	AGC-----CAGATTGCCAAGAACTTCGACGAGAGCCGCTGGCCAAAGAGGTGTCCCT	571							
Db	486	TGCATNNNNNNNNNCGAAGACATTTGAGCAGGAACCCCTAGGAAGGAAGTGTCTTT	545							
Qy	572	GGACGAGGCATCGTGTGTCCTCGCTCCACCGGAGGGCATCCCTCCAGCCGAGGTGGA	631							
Db	546	GGAACGAGGAAGTCTTTACTCCAGTGTGCGACCACTGAAGGGATCCCACTGAGGTGGA	605							
Qy	632	GTGGCTCGGACGAGGACCTGGTGGACCGGTCCCTGGACCCCAATGATATCATCACGG	691							
Db	606	ATGGTTGAAAAAATGAAGACATAATTGATCCCCGTTGAAGATCGGAATTTTATTACTTAT	665							
Qy	692	GGAGCACAGCCTGGTGTGTCGACAGCGCCCGCTTGTCTGACAGCGGCACATCACCTCGCT	751							
Db	666	TGATCACAACTCATATAAGCAGCGCCGACTCTCTGATACTGCAAAATTCACCTGTGT	725							
Qy	752	GGCCAGAAACATCGTGGCAGCTGCGCGGAGCGCCTCCGCTGTGTCTATCGTCTACGTGA	811							
Db	726	TGCCAAAAACATTTGTGCCAAGAGGAAAAAGTACAACTGCCACTGTCTATGTTATGTCAA	785							
Qy	812	CGGTGGGTGGTCGACGTGGAACCGAGTGGTCCGCTCTGCACGCGCCAGCTGTGGCGCGGCTG	871							
Db	786	CGGTGGCTGGTCCACTTGAACGAGTGTGCTGTGTGTAAACAGCGCTGTGGACGAGGTA	845							
Qy	872	GCAGAAACGGACCGAGCTGCAACCAACCCGGCGCCTCTTCAACGGGGGCGCTTTCGTGA	931							
Db	846	TCAGAAACGTACAAGGACTTGTACCAACCCGCGACCACTCAATGGGGGTGCCTTCTGTGA	905							



```
QY 1955 TGCTGCTACGTCTTACCGAGAGCTGGCCGCTTTCCTGCTGGGAGGCGCTCTAG 2014
Db 1971 GGCTGTCACATCTCACAGAACTCAGACCTAGCTGCTGGTAGACATTCACCCAC 2030
QY 2015 CGTGGCTGCGCCAGCGCTCAGCTGCTTCTGTTGGCCGGTGGCCCTGCACTCCCT 2074
Db 2031 CAAAGCGCTGCGAAGCGCTCAGCTGGCCATCTTTGGCCCTGCTGCTGCTCCTGCT 2090
QY 2075 CGAGTACAAATCCGGCTCTACTGCTCCTCATGACACCCAGATGCACTCAAGAGGTGT 2134
Db 2091 GGAGTACAGATCCGAGTCTACTGCTCTGATGACACCCAGATGCCCCGAAGAAATTT 2150
QY 2135 GCAGCTGGAGAGACAGCTGGGGGACAGCTGATCCAGAGCCACGGGTCTCTGCACTCAA 2194
Db 2151 ACATCTTGAGACAGATGGAGGACAGCTCCTAGAAGAACCTTAAGCTCTTCATTITAA 2210
QY 2195 GGAAGCTTACCAACAGCTGGCCCTATCCATCCAGATGTCGCCAGTCCCTGTGGAAGAG 2254
Db 2211 AGGACGACCCCAACCTGGCCCTGCTCAATTCAGATATGCCCATTCCTCTGGAAGAG 2270
QY 2255 TAAGCTCTCTGTGAGCTACAGAGATCCCTTTTATCATCTGGAATGSCAGCGAG 2314
Db 2271 CAAATGCTGGCTAAATATCAGAAATTCATTTTACATGTTGAGTGGATCTCAAAG 2330
QY 2315 GTACTTGGACTGACCTTCACCTGGAGGCTGTGAGCCCGAGCACTAGTGCCTGCTG 2374
Db 2331 AAACCTGCACTGCACCTTCACTCTGAAAGATTTAGCTGAAACACAGTGGAGCTGTTG 2390
QY 2375 CAAAGCTGTGGGTGTCAGGTGGAGGCGAGCGGAGAGCTTCAGATCAACTTCAACAT 2434
Db 2391 CAAACTCTGTGCGGAGGTGAAGAGAGGGGAGATCTTCCAGCTCAACTGCACCGT 2450
QY 2435 CACCAAGGACACAAGGTTTCTGAGCTGCTGCTGAGAGTGAAGCGGGGTCCCGAG 2494
Db 2451 GTCAGAGGAACCTACTGGCATGATTTGCGCTGCTGATCTCTGGAACACCATCACAC 2510
QY 2495 CTTGGTGGCCCGAGTGCCTTCAAGATCCCTTCTCACTTCGGCAGAAATATTTCCAG 2554
Db 2511 GGTCAAGGGGCGGAGTCTTTCAGATCCCTCTCCCTATCCGCGAGAACTCTGTAGCAG 2570
QY 2555 CTTGACCCACCTGTAGCGGGGTGCCACTGGGAGCTGTGGCCAGAACTCCACCT 2614
Db 2571 CTTGATGCCCCCAGACAGAGAGGCACTGAGTGGAGGATGCTGGCCCATAGCTGAACCT 2630
QY 2615 GGACAGCATCTCAGCTTTCTTTCCTCCAAAGCCAGCCCGACAGCCATGATCTCACT 2674
Db 2631 GGACAGGTAATGTAATTTGTCACCAATTCAGCCCACTGCGGATCTCTGGATCT 2690
QY 2675 GTGGAGGGGGGCACTTCCCAACCGCAACCTCAGCCAGCTGGCTGCAGCAGTGGCTGG 2734
Db 2691 TTGGGAAGCAGAGAACTTCCAGATGGAACCTGAGCATGCTGGCAGCTGTCTTGAAGA 2750
QY 2735 ACTGGGCCAGCAGAGCTGGCTCTTACAGTGTGGAGGCTGAGTGTGA 2786
Db 2751 AATGGGAAGCATGAACCGTGGTGTCTTCTTAGCAGCAGAGGGCAGTATTA 2802

RESULT 2
AY406493
LOCUS
DEFINITION Mus musculus UNC5C gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY406493
VERSION AY406493.1 GI:39762467
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2791)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D., and Cargill,M.
Inferred nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 2791)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D., and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
Location/Qualifiers
1..2791
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2791
/gene="UNC5C"
/locus_tag="HCM2575"

ORIGIN
Query Match 31.2%; Score 898.6; DB 29; Length 2791;
Best Local Similarity 61.5%; Pred. No. 1.4e-140;
Matches 1635; Conservative 0; Mismatches 939; Indels 83; Gaps 9;
QY 190 TGCTGGTCCCAACCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACA 249
Db 158 TTCCATCTGACCCACTGAGCCATTGCCACATCTCTCATGAGCCCGAGGAGCTTACA 217
QY 250 TCGTCAAGAACAGCCAGTGTCTTGTGTCAAGCCGCTGCGCCAGCAGATCTTCT 309
Db 218 TTGTGAAGAACAGCCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCT 277
QY 310 TCAAGTGCMAAGGGAGTGGGTGCGCCAGTGGACCACTGATCGAGCGCAGCAGACG 369
Db 278 TCAAGTGCMAAGCGAGTGGGTTCATCAGAAAGCACCGTAGTAGAGAGATAGATG 337
QY 370 GGAGCAGTGGTGGAGCGGACCATGGAGTCCGATTAATGTCTCAAGGCGAGGTCGAGA 429
Db 338 AAACCTCTGTTCTAATGTGAGAGAGTGAAGCAATTCAGATTTACGCCAGCAGTGGAG 397
QY 430 AGGTGTTCCGGCTGGAGGAATACCTGGTCCAGTGGTGGCATGGAGCTCTCGGCACCA 489
Db 398 AACTGTTTGGCCCTGAAGATTACTGTGTCAGTGTGTGGCTGGAGCTCAGCAGGACTTA 457
QY 490 CCAAGAGTCAAGAGCCCTACATCCGCAATAGCCAG-----ATTGCGCAAGAACTTCGAGC 543
Db 458 CGAAGAGTCGAAGAGCATACCTGCGCAATTCAGATCTTCCAGATCTCGGGAAGACATTCGAGC 517
QY 544 AGGAGCCCTGCGCAAGGAGTGTCCCTGGAGCAGGAGCATGCTGCTGCCCTGCCCTCCAC 603
Db 518 AGGACCCCTGGAAAGGAAGTGTCTTGGAGCAGGAGTCTTACTTCCAGTGTTCGGCCAC 577
QY 604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGTGTGACCCGT 663
Db 578 CTGAAGGGATCC-----GTAGAATGGCTAAGAAATGAAGACATAATTTGATCTGT 626
QY 664 CCTTGGACCCCAATGTATACATACGGGGAGCAGACCTGTGGTGGCAGCAGGCCGCC 723
Db 627 CTGAAGATCGGAACCTTTTATATTACTATCGATCAACACCTGATCATCAAGCAAGCCGAC 686
QY 724 TTGCTGACAGCGCCCACTACACCTGCGTGGCAGAAACATGCTGGCAGCTGCCGAGCG 783
Db 687 TCTCAGATACAGCAAAATTTATACCTGTGTGCAAAAAATATTTGTTCCCAAGAAAAAGCA 746
QY 784 CCTCCGCTGTGTCATCGTCTACCTGAACGGTGGTGTGTGACCTGAGCCGAGTGGTCCG 843
Db 747 CCACAGCCACTGTCTATGTTATGTTGCTGCTGCTCCACCTGGACAGAGTGGTGTG 806
```

```
TITLE
JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1..2791
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2791
/gene="UNC5C"
/locus_tag="HCM2575"

ORIGIN
Query Match 31.2%; Score 898.6; DB 29; Length 2791;
Best Local Similarity 61.5%; Pred. No. 1.4e-140;
Matches 1635; Conservative 0; Mismatches 939; Indels 83; Gaps 9;
QY 190 TGCTGGTCCCAACCGGACCTGCTTCCCACTTCTCTGGTGGAGCCGAGGATGTGTACA 249
Db 158 TTCCATCTGACCCACTGAGCCATTGCCACATCTCTCATGAGCCCGAGGAGCTTACA 217
QY 250 TCGTCAAGAACAGCCAGTGTCTTGTGTCAAGCCGCTGCGCCAGCAGATCTTCT 309
Db 218 TTGTGAAGAACAGCCCTGTGAACCTGTATTGTAAAGCCAGCCCTGCCACCCAGATCT 277
QY 310 TCAAGTGCMAAGGGAGTGGGTGCGCCAGTGGACCACTGATCGAGCGCAGCAGACG 369
Db 278 TCAAGTGCMAAGCGAGTGGGTTCATCAGAAAGCACCGTAGTAGAGAGATAGATG 337
QY 370 GGAGCAGTGGTGGAGCGGACCATGGAGTCCGATTAATGTCTCAAGGCGAGGTCGAGA 429
Db 338 AAACCTCTGTTCTAATGTGAGAGAGTGAAGCAATTCAGATTTACGCCAGCAGTGGAG 397
QY 430 AGGTGTTCCGGCTGGAGGAATACCTGGTCCAGTGGTGGCATGGAGCTCTCGGCACCA 489
Db 398 AACTGTTTGGCCCTGAAGATTACTGTGTCAGTGTGTGGCTGGAGCTCAGCAGGACTTA 457
QY 490 CCAAGAGTCAAGAGCCCTACATCCGCAATAGCCAG-----ATTGCGCAAGAACTTCGAGC 543
Db 458 CGAAGAGTCGAAGAGCATACCTGCGCAATTCAGATCTTCCAGATCTCGGGAAGACATTCGAGC 517
QY 544 AGGAGCCCTGCGCAAGGAGTGTCCCTGGAGCAGGAGCATGCTGCTGCCCTGCCCTCCAC 603
Db 518 AGGACCCCTGGAAAGGAAGTGTCTTGGAGCAGGAGTCTTACTTCCAGTGTTCGGCCAC 577
QY 604 CGGAGGGCATCCCTCCAGCCGAGGTGGAGTGGCTCCGGAACGAGGACCTGTGTGACCCGT 663
Db 578 CTGAAGGGATCC-----GTAGAATGGCTAAGAAATGAAGACATAATTTGATCTGT 626
QY 664 CCTTGGACCCCAATGTATACATACGGGGAGCAGACCTGTGGTGGCAGCAGGCCGCC 723
Db 627 CTGAAGATCGGAACCTTTTATATTACTATCGATCAACACCTGATCATCAAGCAAGCCGAC 686
QY 724 TTGCTGACAGCGCCCACTACACCTGCGTGGCAGAAACATGCTGGCAGCTGCCGAGCG 783
Db 687 TCTCAGATACAGCAAAATTTATACCTGTGTGCAAAAAATATTTGTTCCCAAGAAAAAGCA 746
QY 784 CCTCCGCTGTGTCATCGTCTACCTGAACGGTGGTGTGTGACCTGAGCCGAGTGGTCCG 843
Db 747 CCACAGCCACTGTCTATGTTATGTTGCTGCTGCTCCACCTGGACAGAGTGGTGTG 806
```

QY 844 TCTGAGGCGCAGCTGTGGGCGCGCTGGCAGAACGAGCGGAGCTGCACCAACCGG 903  
DB 807 TGTGTAAACAGCGCTGTGGGCGAGGATATCAGAAACGCACAGAACCTGCAACCCAG 866  
QY 904 CGCTCTCAACGGGGCGCTTCTGTGAGGGGCAGAAATGTCATACCGCGACCGTCTCCT 963  
DB 867 CCCACTCAATGGTGGGCGCTTCTGTGAGGGGCAGAGTGTGCAGAA---AATAGATGCA 923  
QY 964 CTCTGCTTGTCTGTGTGACGGCAGCTGGAGCCCGTGGAGAAAGTGTGGCGCTGTGGG 1023  
DB 924 CTACGTTATGTCACGTGATGTAGTGTGCTCTCATGGAGCAATGTCTCAACTGTGGGA 983  
QY 1024 TGGACTGACCCACTGGGCGAGCGTGTGCTCTGACCCAGCACCCCGCAACGAGGG 1083  
DB 984 CTGAATGACCCACTGGGCGAGGAGGTGTACAGCACCCCGCAAGAACGGGGTA 1043  
QY 1084 AGGAGTGCACGGGCACTGACCTGGGACACCCGCAACTGTACCAAGTGACCTCTGTATACCA 1143  
DB 1044 AGGAGTGTGATGGCTGTGCTTCCATCAAGAACTGCACTATGGGCTGTGCATGCAGG 1103  
QY 1144 GTGCTTCTGGCCCTGAGGAGCGTGTCTATGTGG---CCTCATGGCGGTGGCGCT 1200  
DB 1104 CTGCTCTGACTCAGATGATGTGGCTCTCTACGTGGGATTTGTGATCGCTGAACAGTCT 1163  
QY 1201 GCCTGGTCTGTGCTGTCTCTCATCTCTGTTTATGCGGAGAGGAGGGGCTGG 1260  
DB 1164 STCTGGCGATCACTGTGTGGTGGGCGCTGTTGTGTATCGGAAGAACCCGTGACTTTG 1223  
QY 1261 ACTCAGATGTGGCTGACTGCTGCATCTCACTCTCAGGCTTCAGGCGCGTCAAGCTCAAGC 1320  
DB 1224 AGTCTGACATCATGACTCTCAGACTCAATGGGCGCTTCAGGCTGTGAACATCAAGG 1283  
QY 1321 CCAGCAAGCAGACAAACCCCAATCTGCTCAACATCAGCGGAGCTTCAGCAACACCA 1380  
DB 1284 CTGCAGACAAGATCTCTGGTGTCTCCCGCTGACC-----TCACCTCAGCTGCAGCA 1337  
QY 1381 CCTACAGGAGTCTGTCTCCCGCAGAGT-----GGCCGAGCCCAAGTTCAGCT 1435  
DB 1338 TGTACAGGGGACCTGTGTATGCTGTGATGCTCTCAGACAAATCTCCCAATGACCACT 1397  
QY 1436 CACCAATGGGCACTCTCAGCGCCCTGGGTGGCGCCGCCACACACTGCAACACAGCTC 1495  
DB 1398 CTCCAATCTGACCACTACCACTTGAATCAAGTGTACACAGCTCAGGTGCTG 1457  
QY 1496 TCCCACTC-----TGAGGCGGAGGAGTGTCTCTCCCGCTCTCCACC----- 1539  
DB 1458 TCACTCTCAGGATGACCTTGGCGAGTCTCTCATCCAACTGTCTCACCAGATGACCCAGT 1517  
QY 1540 -----AGAACTACTTCCGCTCCCTGCGGAGGCA 1569  
DB 1518 CTTGTGTAGAGATGAGGCGCTTAACCTGAAGAACAGAGCTTCGGAAGACAGCTGACC 1577  
QY 1570 CCAGCAACATGACTATGGGACCTTCAACTCTCTCGGGGCGCGGCTGATGATCCCTAATA 1629  
DB 1578 CATCTGTGACAGCAATTTGTAACCTTCACTCTCTTGGGGTCACTCATCTTCTCAAT 1637  
QY 1630 CAGGTATGACCTCTCATCTCCCGCAGATGCCATACCCGAGGAGATCTATGATATCT 1689  
DB 1638 CAGGAGTAAGCTTGTCTGATTCGCGTGGGCGCAATCTCTCAGGGGAGAGTCTATGAATGT 1697  
QY 1690 ACTCAGCTGACAGCGGAGAGCTGAGGTGTGCGCTTGCCTTGTGCTGCTGACAGCTGTC 1749  
DB 1698 ATGTGACTGTACACAGAAAGAAATATGAGGCGCCCGCATGGAAGACTCTCAGACCTTAC 1757  
QY 1750 TGAGTCCCATGCTGTAGTGTGAGCCCGCTGGCGTCTGTGCTCACCGGCGCAGTCACTCTGG 1809  
DB 1758 TTACCCCTGTGGTGTGAGTGTGGGCTCTCTGAGCTCTGTGCTGACCGCGCTGTCTATCTCA 1817  
QY 1810 CTATGACACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGCGCTCAAAAGCAGT 1869  
DB 1818 CTCTGATCATCTGTGACAGCCCGCAGCACCGAGGACTGGAGATCCAGCTCAAAAACGAG 1877  
QY 1870 CGTGGAGGGCAGCTGGGAGCAGATGTGCTGCACTGGGCGAGGAGGCGCCCTCCACC 1929

DB 1878 CAGTGCAGGACAAATGGGAGGATGTTGTGTG---GTTGGGAGGAGAACTTCAAAACC 1934  
QY 1930 TCTACTACTGCGAGCTGGAGCGCAGTGCCTCTCTAGTCTTCTCACGAGCAGTGGCGCGT 1989  
DB 1935 CTTGTATCAATCAGCTGTGATGACAGGGCTTCCCATATCTCTCACAGAACTCTAGTACT 1994  
QY 1990 TTGCGCTGTGGGAGAGCCCTCAGCGTGGTGCGCCCAAGCGCTCAAGCTGCTTCTGT 2049  
DB 1995 ATGCCCTGTGTTGGGCACTGCCACCAAGCAGCTGCCAAGGCTCTTAACTGGCCATCT 2054  
QY 2050 TTGCGCGCGTGGCTGACCTCCCTCGAGTACAAATCCGGGTCTACTGCTGTGATGACA 2109  
DB 2055 TTGCGCGCTCTGTGCTCTTCCCTGGAGTACAGCAATTAGCTCTACTGCTGTGATGACA 2114  
QY 2110 CCACAGTCACTCAAGGAGTGTGAGCTGGAGAACAGCTGGGGGAGACAGTATCC 2169  
DB 2115 CACAGATGCCCTGAAGGAAGTTCTACAACTGGAGAGGCAATGGAGGAGACAGCTCTTAG 2174  
QY 2170 AGGAGCCAGGCTCTGCACTTCAAGGACAGTTTACCAAACTGGCGCTTATCATCCAG 2229  
DB 2175 AAGAACCAGGCTCTTCAATTAAGGAGCAGCATCCACAACTGCGCTGTCTATTATG 2234  
QY 2230 ATGTGCCAGCTCCCTGTGGAAGATAAGCTCTCTGTGAGCTACAGAGATCCCTTTT 2289  
DB 2235 ACATCGCCCATTCCTCTGGAAGAGCAAAATGCTGGTAAAGTATCAGGAAATTCATTTT 2294  
QY 2290 ATCACTGTGNAATGGCAGCGAGCGGTACTTGCCTGCACTGCACTTCACTTCAAGTGGAGCGTCA 2349  
DB 2295 ACCACATCTGGAGTGGCTCTCAAGAAACCTCCCTGCACTTCACTTGGAAAGACTCA 2354  
QY 2350 GCCCAGCACTAGTCACTGGCTGCAAGCTGTGGGTGTGGAGGTGGAGGCGAGCGGC 2409  
DB 2355 GCTTAAACACAGTGTAACTGTTTGCMAACTCTGTGTGGCGCAGGTTGAAGGAGAGGCG 2414  
QY 2410 AGAGCTTCAGATCAACTTCAACATCACCAAGACACAAAGGTTTGTGAGCTGTGCTGCTC 2469  
DB 2415 AGATCTTCCAGCTCAACTGTACTGTGTAGAGGAACTTACTGGCATCGACTTACTCTCC 2474  
QY 2470 TGGAGATGAAGCGGGGTCCAGCCCTGTGGGCGCCAGTGCCTTCAAGATCCCTTCC 2529  
DB 2475 TGGACCTCTAGTACCATCACTGTCACCGGACCAAGTGTCTTTCAGANTTCTCTCC 2534  
QY 2530 TCATTTCGGCAGAGATAATTTCCAGCTCGAACCCCTGTAGTGGCGGGTGGCGACTGGC 2589  
DB 2535 CTATCGGCGAAGCTATGCAAGCGCTGATGCCCTCAAACAAGAGGCCATGACTGGA 2594  
QY 2590 GGACTTGGCCCGAGAACTCCACCTGGGACAGCCATCTCAGCTTCTTGGCTCCAAAGCCCA 2649  
DB 2595 GGATGTGGCCCATAACTCAACCTGGACAGGTACTTGAATTACTTTGGCCACCAATCGA 2654  
QY 2650 GCCCAGACGCTATCTCACTGTTGGGAGCGCGGCACTTCCCAACGGCAACTCA 2709  
DB 2655 GCCCACTGGCTAACTCTCTGGATCTTTGGGAAGACAGAACTTCCAGATGGAACCTGA 2714  
QY 2710 GCCAGCTGTGTCAGCAGTGGCTGGACTGGGCGAGCCAGACGCTGGCTCTTTCACAGTGT 2769  
DB 2715 GCATGTGGCAGCGCTCTCTGGAGAAATGGGAAGACATGAGACAGTGGTGTCTTGGCAG 2774  
QY 2770 CGGAGCTGAGTCTGA 2786  
DB 2775 CAGAAGGACAGTATTGA 2791

## RESULT 3

AK031655

LOCUS

DEFINITION

AK031655 3790 bp mRNA linear HTC 18-SEP-2003  
Mus musculus 13 days embryo male testis cDNA, RIKEN full-length  
enriched library, clone:6030473H24 product:unc5 homolog (C.  
elegans) 3, full insert sequence.

ACCESSION

AK031655

VERSION

AK031655.1 GI:26327502

KEYWORDS

HTC; CAP trapper.

SOURCE	Mus musculus (house mouse)																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	</
--------	----------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

QY 670 ACCCAATGATATACATACAGCGGGAGGACAGCGTGTGGTGGAGAGCGCGCTTGCTG 729  
DB 755 ATCGAACTTTTATATTAATCATCATCAACCTGATCATCAAGCAAGCCCGACTCTCAG 814  
QY 730 ACAGCGCAACTACACCTCGTGGCCAAAGAACATCGTGGCAGCTCGCGCAGCGCTCGG 789  
DB 815 ATACAGCAAAATTAACCTGTGTGGCCAAATATTTGGCCAGAGAAAAGCACACAG 874  
QY 790 CTGTGTCTCATCGTCTACGTGAACGCTGGGTGTGTCAGCTGGACCGAGTGGTCCGCTGCA 849  
DB 875 CCACGTGTATGTTAATGTGTGCTGTCCACCTGGACAGAGTGTGTGTGTA 934  
QY 850 GGGCAGCTGTGGCGGCGCTGGCAGAAACGAGCGGAGCTGCACCAACCGCGCGCTC 909  
DB 935 ACAGCGCTGTGGCGGAGGATATCAAAACGCAAGAACCTGCACCAACCGAGCGCCAC 994  
QY 910 TCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCATGACCGCACCGCTCTCTCTGC 969  
DB 995 TCAATGTGGGCGCTTCTGTGAGGGGAGAGTGTGAGAA---AATAGATGACTACGT 1051  
QY 970 TTGTCTGTGGACGGCAGCTGGAGCGCGCTGGAGCAAGTGGTGGCGCTGTGGGCTGGACT 1029  
DB 1052 TATGTCCAGTGTGATGTAGTGGACTTTCATGAGCAAAATGCTCAACCTGTGGGACTGAAT 1111  
QY 1030 GCACCCACTGGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGAGGGGAGGT 1089  
DB 1112 GCACCCACTGGCGGAGGGAGTGTACAGCACCGACCGCCCAAGAACCGGGGTAAAGACT 1171  
QY 1090 GCCAGGGACTGACCTGGACACCCGCACTGTACAGTGAACCTCTGTGTACACAG---- 1144  
DB 1172 GTATGGCGCTGTCTTCCAAATCAAGAACTGCACTGATGGGCTGTGATGAGGGATTCA 1231  
QY 1145 -----TGCTTCTG 1152  
DB 1232 TTTTACCCATTTCACTGAGCACAGACCCAGAAATGAATATGGAATTTTCTGTCTCTG 1291  
QY 1153 GCCTGTGAGACGTGGCGCTCTATGTGG---CCTCATCGCGTGGCGCTGCTCGTCTCC 1209  
DB 1292 ACTCAGATGATGGCTCTCTACGTGGGATTTGTATCGCTGTAAACGCTGTCTGGGCA 1351  
QY 1210 TGTGTCTGTGTCTCTCATCTCTGTTTATGTCGGGAAGAGGGGCTGGACTCAGATG 1269  
DB 1352 TCACTGTGTGTGGCGCTCTGTGTGTATCGGAAGAACCCCGTGACTTTGAGTCTGACA 1411  
QY 1270 TGGCTGACTGTGCTCACTCTCACTCAGGTTCAGCCCGTTCAGCATCAAGCCAGCAAG 1329  
DB 1412 TCATTGACTCTCAGCACTCAATGGCGGCTTTTCAGCCTGTGAACATCAAGGCTGCCAGAC 1471  
QY 1330 CAGACAAACCCCATCTGCTCAACATCCAGCGGACCTCAGCACCAACCACTACAGG 1389  
DB 1472 AAGATCTCTGTGTGTCT 1525  
QY 1390 GCAGTCTCTGTCT 1444  
DB 1526 GACCTGTCTATGCTCTGATGATGTCTCAGACAAATCCCAATGACCACTCTCCAATC 1585  
QY 1445 GCACCTGTCTAGCCCTCTGGTGGCGGCGCCACACATGTGCACACAGCTCTCCAGCTC 1504  
DB 1586 TGGACCACTACCCAACTTGAAATCAAAAGTGTACAAACAGCTCAGGTGTGTCTCTCT 1645  
QY 1505 -----TCAGGCGGAGGAGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1539  
DB 1646 AGGATGACCTTGGAGTGTCT 1705  
QY 1540 -----AGAACTACTTTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1578  
DB 1706 AGAATGAGGCGCTTAACCTGAAGAAACAGAGCTCGCAAGACAGACTGACCCATCTGCA 1765  
QY 1579 TGACCTTANGGACCTTCACT 1638  
DB 1766 CAGCATTTGTACCTTCACT 1825  
QY 1639 GCCT 1698

DB 1826 GCTTCTGATTTCCCGCTGGGSCCATTCCTCAGGGAGAGTCTATGAAATGATGTGACTG 1885  
QY 1699 TGCACAGCCGGAAGACGTGAGGTTGGCCCTAGCTGGTGTGAGACCCCTGCTGAGTCCCA 1758  
DB 1886 TACACAGAAAGAAAATATGAGGCCCCCCTCAGAAAGACTCTCAGACCCCTACTTACCCCTG 1945  
QY 1759 TCGTTAGCTGTGACCCCTCGCTGCTGCTCACCCGCCAGTCTATCTGGGTATGGACC 1818  
DB 1946 TGGTGTGCTGTGGGCTCTCTGAGCTCTGCTGACCCGCTGTCTCTCTCTCTCTCTCTCTCT 2005  
QY 1819 ACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGCGCTCAAAAGAGAGTGTGCGAGG 1878  
DB 2006 ACTGTGACAGCCAGACCCAGGAGTGGAAATCCAGCTCAAAAACAGGAGTGCAGG 2065  
QY 1879 GCAGCTGGAGAGAGATGTGTGACCTGGGGGAGGAGGCGCTCCACCTCTACTACT 1938  
DB 2066 GACAAATGGAGGATTTGTGGTG---GTGGGGAGGAGAACTTCAACCCCTCTTTACA 2122  
QY 1939 GCACGTGGAGCCAGTGTCTGTCTTACCGAGCAGCTGGGCGCTTTTGGCCTGG 1998  
DB 2123 TTCACTGTGATGACAGGCTTGGCATATCTTACAGAGAACTCAGTACCTATGCTCTGG 2182  
QY 1999 TGGGAGAGCCCTCAGCGTGGCTGCGCCAAAGCGCTCAAGCTGCTCTGTGTTGCGCGG 2058  
DB 2183 TTGGGAGTCCACCAACCAAGCAGCTGCCAAGCGCTTAAACCTGGCCATCTTTGGGCCCC 2242  
QY 2059 TGGCTGTGACCTCCCTCGAGTACAACTCCGGTCTACTGCGCTGATGACACCCAGGANG 2118  
DB 2243 TCTGTGTCT 2302  
QY 2119 CACTCAAGGAGTGTGTGAGTGGAGAGCAGCTGGGGGAGCAGCTGTATCCAGGAGCCAC 2178  
DB 2303 CCTCAAGGAGTCTCACTGTGGAGGCAATGGAGGACAGCTCTTAGAAGACCCA 2362  
QY 2179 GGGTCTCTGACTTCAAGGACAGTTACCAAACTGGCGCTTATCCATCCAGTGTGCCCA 2238  
DB 2363 AGGCTCTTCTGTTTAAAGGAGCAGTCCCAAACTGGCGCTCTTATTCATGACATCGCCC 2422  
QY 2239 GCTCCCTGTGGAAGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2298  
DB 2423 ATTCCCTCTGGAAGCAAAATGCTGGTAAAGTATCAGGAAATTCATTTTACCACTCT 2482  
QY 2299 GAAATGGCAGCAGCGGTACTTGCATGCACTTCACTTCACTTCACTTCACTTCACTTCACT 2358  
DB 2483 GAGTGTCTCTCAAGAACTCTCACTGCACTTCACTTCACTTCACTTCACTTCACTTCACT 2542  
QY 2359 CTAGTACCTGGCTGCAAGCTGTGGTGTGGAGGTGGAGGGCGACGGGAGAGCTTCA 2418  
DB 2543 CAGTGAACCTGTTTGCAAACTCTGTGTGGCAGGTTGAAGGAGAGGAGAGATCTTCC 2502  
QY 2419 GATCAACTTCAACATCAACAGGACCAAGGTTTGTGCTGAGCTGCTGCTGCTGAGAGTG 2478  
DB 2603 AGCTCAACTGTACTGTGTGAGAGAACCTTCTGAGCACTTACCTCTCTCTGAGACCTG 2562  
QY 2479 AAGCGGGGTCCAGCGCTTGGGGCCCGAGTGGCTTCAAGATCCCTCTCTCTCTCTCTCT 2538  
DB 2663 CTAGTACCATCACTGTGTGACCGAACAGTGTCTTCACTTCTCTCTCTCTCTCTCTCT 2722  
QY 2539 AGAAGATAATTTCCAGCTGGACCCACCTGTAGGGGGGTGCGAGTGGGAGCTCTGG 2598  
DB 2723 AGAAGTATGACGACCTGGATGGCCCTCAAAAGAGGCGCATGCTGAGGATGCTGG 2782  
QY 2599 CCCAGAACTCCACCTGGACAGCATCTCAGCTTCTTGTCTTCAAGCCAGCCCGACAG 2658  
DB 2783 CCATAACTCAACCTGGACAGTACTTGAATTTCTTTGCGCAAAATCAGGCCAATG 2842  
QY 2659 CCATGATCTCAACCTGTGGAGGCGGCGACTTCCCAACGGCAACCTCAGCGAGCTG 2718  
DB 2843 GCGTATCTGTGATCTTTGGAGAGCAGAACTTCCAGATGAAACCTCAGCATGTGG 2902  
QY 2719 CTGAGAGTGTGCTGAGCTGGGCGCAGCAGACCTGCGCTCTTTCACAGTGTGAGGCTG 2778



Db 600 AACAGCCAGTGGAACTGCACTGAGAGCCCTTCCCTGCCACGCACTACTTCAAGTGT 659  
Qy 318 AACGGGAGTGGGTGCGCCAGGTGACCACTGATCGAGCGCAGACAGACGGGAGAGT 377  
Db 660 AATGSCGAGTGGGTGAGCCAGAAAGACACAGTCAACAGGAGAGCTGGATGAGGCCACA 719  
Qy 378 GGTGAGCCAGCATGAGAGTCCGATTAAATGCTCAAGGAGAGAGTGGAGAGTGTTC 437  
Db 720 GGTGTGCGGGTCCGAGAGGTGAGATCGAGGTGTCAAGGAGAGTGGAGAACTCTTC 779  
Qy 438 GGGCTGAGGAGAACTACTGCTGCGAGTGGATGGAGCTCTCGGGCACCACCAAGAGT 497  
Db 780 GGGCTGAGGAGTACTGCTGCGAGTGGATGGAGCTCTCGGGCACCACCAAGAGT 839  
Qy 498 CAGAGGCTTACATCCGATAGCCAGATGTCGCAAGAACTTCGAGAGAGAGCCGTGGCC 557  
Db 840 CGCCGAGCTTACATCCGATGCTTCTGCGCAAGAACTTCGAGAGAGCTTCGGCC 899  
Qy 558 AAGGAGGTGTCCCTGAGCAGGAGCATCGTCTGCCCTCGCTCCACCGGAGGGCATCCCT 617  
Db 900 AAGGAGGTACCTTGGATCATGAGTCTTCTGCAAGTGGCGCCACCGGAGGAGTGCCT 959  
Qy 618 CCAGCCGAGGTGAGTGGCTCGGAAACGAGGACCTGGTGGACCCCTCCCTGGACCCCAAT 677  
Db 960 GTGGCTGAGGTGGAATGGCTCAAGATGAAGATGCAATTGACCCCGCTCAGGACCTAAC 1019  
Qy 678 GTATACATACCGGGAGCACAGCTGGTGGTGGACAGGCGCGCTTCTGACAGGCG 737  
Db 1020 TTCTGCTCACCATTGACCAACCTCATATCCCGCAGGCGCGCTCTCAGACAGGCG 1079  
Qy 738 AACTACCTCGTGGCCAGAAACATCGTGGACGTCGCGCAGCGCTCCGCTGGTGTTC 797  
Db 1080 AACTACCTGTGTGGCCAGAAATATGTTGGCCAGCGCGGAGACCGCGGCCACAGTC 1139  
Qy 798 ATCGTCTACGTGAACGGTGGGTGGTGCAGCTGGACCGAGTGGTCCGTCTGACGCGCCAGC 857  
Db 1140 ATCGTCTATGTGAATGGAGGCTGGTCCAGCTGGCAGAGTGGTCACTCTGTTCCAAATGC 1199  
Qy 858 TGTGGCGCGGTGGCAGAAACGAGCCGAGTGGACCAACCCGCGGCTCTCAACGGG 917  
Db 1200 TGTGGCCGAGGTGGCAGAAAGCTACTCGGACCTGCACCAATCCAGCCCGCTCAATGGA 1259  
Qy 918 GCGCTTTCTGTGAGGGCAGAAATGTCATGACCGCACCGTCTCTCTGCTGTGCTCT 977  
Db 1260 GCGCTTTCTGTGAGGGACAGGCTTCCAGAGACAGCTTGCACACCGTGTG---CCCA 1316  
Qy 978 GTGGACGCACTGGAGCCCGTGGAGCAAGTGGTCCGCTGGTGGGTGGAGTGCACCCAC 1037  
Db 1317 GTGGATGGAGCGTGGACCGAGTGGAGCAAGTGGTCTGCTGACAGCAGAGTGTGGCCAC 1376  
Qy 1038 TGGCGGAGCGGTGAGTGTCTGACCCAGCACCCCGCAACGGAGGGAGGAGTCCAGGGC 1097  
Db 1377 TGGCGGAGCGCGAGTGCATGGCACCGCCACCCCAAGAAACGGAGCGCGTGACTCAGCGGG 1436  
Qy 1098 ACTGACCTGACACCCGCAACTGTACAGTGAACCTGTGTGTACACAGTCTTCTGGCC-- 1155  
Db 1437 AGCTACTTGTATCCAGAACTGCACTGATGGGCTGTGGTGTGTGATCAGAGAACTCTA 1496  
Qy 1156 -----CTGAGGAGTGGCCCTCTATGTGGGCTC 1184  
Db 1497 AACGACCTTAAAGCCACCCCTGGAGACATCGGAGATGGCACTGTACGAGGCGCTT 1556  
Qy 1185 ATGCGCGTGGCCCTGCGCTGGTCTGCTGCTGCTGTC---CTCATCTCGTTTATGTC 1241  
Db 1557 GTGGTGGCGCTCTTGTGGTGGTGGCGGTCTCATGGCCGAGGAGTGTATGATACCGG 1616  
Qy 1242 CGGAGAGAGGAGGGGTGGAATCAGATGTGGCTGACTCGT---CCATTCCTCACTCAGGC 1298  
Db 1617 AGAACTGCGGAGCTTCGACACGACATCACCGACTCTCTGCGGCCCTCACTGGTGGC 1676  
Qy 1299 TTCCAGCCCTGATGATCAAGCCAGCAAGAGCAGCAACCCCTATCTGCT-----CACC 1352  
Db 1677 TTCCACCTGTCACTTCAAGCTGCAAGGCGCCCAACACCCGCGAGCTCTCTGCACCCGCTCC 1736

Qy 1353 ATCCAGCCGAGCTCAGCACCAACCACTTACAGGGCAGTCTCTGTCTCCCGGAGGAT 1412  
Db 1737 GCCCTCAGACCTTAAACGGCCAGTGTGGCATCTACCGCGGCGCTGTGTATGCCCTGCAG 1796  
Qy 1413 GGGCCAGCCCAAGTTCAGCTCAACATGGGCACTGTCT----- 1453  
Db 1797 GACTCCGCGCAAGATCCCATGACTAATTCGCGCCCTGCTGGATCCCTTCCCGCAGCTC 1856  
Qy 1454 -----C 1454  
Db 1857 AAGATCAAGGTCTATACTCCAGCACCAATCGGTTCTGGGTCTGGCTGGCTGATGAGCC 1916  
Qy 1455 AGCCCTCGGTGGCGCGCCACACACTGACCACTCTCCACCTCTTGAAGCCGAG 1514  
Db 1917 GACTGTCTGGGTGTCTCCCGCGGCGAGTACCCAGGGGATTTCTCCCGGACACCCAT 1976  
Qy 1515 GAGTCTGTCTCCCGCTCTCCACCCAG-----AACTATTCTCGGTCTCCCTGCCCGGAGC 1568  
Db 1977 TTCTGCACTGCGGAGTCCAGCTTGGTTTCCAGCACCTCTCTGGGCTTACTCCGGAC 2036  
Qy 1569 ACCAGCAATGACCTATGGGACCTTCAACTTCTCCGGGGCGCTGATGATCCCTAAT 1628  
Db 2037 CCCAGCAGAGTGTGAGCGGACCTTGGTTCCCTGGAGGAGGCTGAGCCTCCCGGC 2096  
Qy 1629 ACAGGTATCAGCTCTCTCATCCCGCAGATGCCATACCCCGAGGGAAGATCTATGATC 1688  
Db 2097 ACAGGGTCTGAGCTGTGGTACCAATGAGGCCATTTCCCGAGGCAAGTTCTATGACCTG 2156  
Qy 1689 TACTCAGCTGCAACAGCGGAGAGCTGAGTGGTGGCCCTAGCTGGCTGTGACACCTG 1748  
Db 2157 TATCTACATATCAACAAAGCCGAAAGCACCTTCCCAAGTTTCCAGAGTTTCCAGACAGTA 2216  
Qy 1749 CTGAGTCCCATCTGTAGCTGTGGACCCCTCGCTGCTCTCACCGGCGAGTCACTCTG 1808  
Db 2217 TTGAGCCCTCGGTGACCTGTGGGCCACAGGCTTACTCTGTGCGGCCCTGTCTCTC 2276  
Qy 1809 GCTATGGACCACTGTGGGAGCCCGCTGACGTGAGCTGCGCTCGCTCAAAAAGCAG 1868  
Db 2277 ACGTGGCCCACTGTGCTGAAGTATCGCTGGAGACTGATCTTTTCAAGACCCAG 2336  
Qy 1869 TGTGGAGGAGCTGGAGCAGAGTGTGCTGCACCTGGCGGAGGAGGCGCCCTCCAC 1928  
Db 2337 GCGCATCAGGCGCACTGGG---AGGAGTGGTGAACCTTGGATGAGGAGACCTTCAACACA 2393  
Qy 1929 CTCTACTCTGCAGCTGGAGCCAGTGTCTGTCTTACCGAGAGCTGGGCGCGC 1988  
Db 2394 CCTGTCTGCTGAGCTGGAGGCTAAGTCTTGGCCATCTCTGTGACAGCTGGGTTC 2453  
Qy 1989 TTTGCCCTGTGGGAGGCGCTTCTGAGCTGGCTGGCGCAAGCGCTCAAGCTGTCTG 2048  
Db 2454 TACGTATTCTGCGGAGTCTTCTCTGTCTGAGTCAAGCGGCTCCAGCTGGCCATC 2513  
Qy 2049 TTTGGCCGCTGCGACCTCTCGAGTCAACATCCGGTCTACTCGCTGCTGCTGATGAC 2108  
Db 2514 TTGCGCCAGCCCTCTGCACTCTCCCTGGAGTAGCTCAGGCTCAGGCTTACTGTCTGGAGGAC 2573  
Qy 2109 ACCCAGATGCACTCAAGAGGTGTGAGCTGGAGAGCAGCTGGGGGAGACAGCTGATC 2168  
Db 2574 ACNCTGTAGACTGAAGAGGCTCTGGAGCTGGAGGAGTCTTGGTGGCTACTTGGTG 2633  
Qy 2169 CAGGAGCCCGGTCTGCACTTCAAGGACAGTTTCAACAACTGGCGCTTATCCATCCAC 2228  
Db 2634 GAGGAGCCCAAGCTTGTCTTTAAGGACAGTTACCAACCTAGCGCTCTCCCTCCAT 2693  
Qy 2229 GATGTGCCAGCTCCCTGTGGAGAGTGAAGTCTTGTGAGCTACAGGAGATCCCTTT 2288  
Db 2694 GACATCCCCCATGCCACCTGGAGGCAAACTACTGGCCAAAGTACAGGAGATTCCTTC 2753  
Qy 2289 TATCATCTCGAAATGGCAGCGAGCTACTTGTCACTGCACTTCACTTGGAGCGTGTG 2348  
Db 2754 TACCACTGTGAATGGCAGCGAGAGCCCTGCACCTGCACTTTTACCTTGGAGAGCAT 2813



QY 2349 AGCCCGACACTAGTGAACCTGGCTCCTAAGCTGTGGTGTGGCAGGTGAGGGCGACGG 2408  
 Db AGCTTGGCTTCCACGAGTTCACCTGAAGTCTGGTGGCAGGTGCAAGGGGAGGC 2873  
 QY 2409 CAGAGCTTCAGCATCACTTCAACATCAC--AAGGACACAAGTTTCTGAGCTGCTG 2465  
 Db CAGATTTTCCAGCTGCACACAAGTTTGGCGGAGACGCTGCTGCTCCCTGGATGCTCTC 2933  
 QY 2466 GCTCTGGAGAGTGAAGCGGGGTCCAGCCCTGGTGGGCGCCAGTGCCTTCAGATCCCC 2525  
 Db TGCTCTGCCCCGGGCGATGCCATCACACCCAGCTGGGACCCCTATGCCCTTCAAGATACCC 2993  
 QY 2526 TTCCTCAATTCGGCAGAGAATAATTTCCAGCTGGACCCACCTGTAGGCGGGTGCAGAC 2585  
 Db CTGTCCATCCGCCAAAGATCTGCAGCAGCTGGAGCCCGCCGACTCCCGGGGCAACGAC 3053  
 QY 2586 TGGCGACTCTGGCCCGAGAACTCCACTTGGACAGCATCTCAGCTTCTTTTGCCTCCAAAG 2645  
 Db TGGAGGCTGTTGGCGCAGAAGTGTCCATGACCGGTACTAACTACTTTCGCCACCAA 3113  
 QY 2646 CCCAGCCCCACAGCCATGATCTCAACCTGTGGGAGGCGGCGACTTCCCCAACGGCAAC 2705  
 Db GCTAGTCCACAGAGTGTCACTTAGACCTCTGGAGAGCTTCGGCAAGATGACGGGGAC 3173  
 QY 2706 CTCAGCCAGCTGGTGCAGCAGTGGCTGAGCTGGGCGACCCAGACGCTGGCTCTTCA 2765  
 Db CTCAACAGCTGGCCAGTGCCTTGGAGGAGATGGGCAAGAGTGAGATGCTGTAGCCATG 3233  
 QY 2766 GTGTCGGAGGCTGAGTGTCTGAG 2787  
 Db GCCACAGATGGCGATTGCTGAG 3255

RESULT 5  
 LOCUS AY411747  
 DEFINITION Homo sapiens HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.  
 ACCESSION AY411747  
 VERSION AY411747.1 GI:39767715  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 2532)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
 Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
 Science 302 (5652), 1960-1963 (2003)  
 14671302  
 2 (bases 1 to 2532)  
 Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.  
 Direct Submission  
 Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
 This sequence was made by sequencing genomic exons and ordering them based on alignment.  
 FEATURES  
 source  
 1..2532  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 <1..>2532  
 /locus\_tag="HCM4327"  
 gene  
 ORIGIN

Query Match 26.5%; Score 763.2; DB 29; Length 2532;  
 Best Local Similarity 60.0%; Pred. No. 7.6e-118;  
 Matches 1515; Conservative 0; Mismatches 873; Indels 138; Gaps 9;  
 QY 393 GAGGTCCGCTTAAATGCTCTCAAGGCAGCAGGTCTGAGAAGTGTTCGGGCTCGAGGAATAC 452  
 Db 13 GAGTGCAGATCGAGGTGTTCGGCGCAGCAGGTGGAGAGCTCTTTGGGCTGGAGATTAC 72  
 QY 453 TGGTGCAGTGTGGTGGCATGAGAGCTCTCGGACACCAACAGAGTCAGAAAGGCTTACATC 512  
 Db 73 TGGTGCAGTGTGGTGGCATGAGAGCTCTCGGACACCAACAGAGTCAGAAAGGCTTACATC 132  
 QY 513 CGCATAGCCAGATTGGGCAAGAACTTCAGCAGGAGAGCGCTGGCCAGAGAGTGTCCCTG 572  
 Db 133 CGCATAGCCCTACTGCGCAAGAACTTCGATCAGGAGCCTCTGGGCAAGAGGTGCCCTG 192  
 QY 573 GAGCAGGGCATCTGTCTGCCCTCCCGTCCACCGAGGGCATCCCTCCAGCGAGGTGGAG 632  
 Db 193 GACCATGAGTTCTCTCTGAGTCCCGCGCGAGGGGTGCTGTGGCCGAGGTGGAA 252  
 QY 633 TGGCTCCGAAAGAGAGACTTGGTGAACCGCTCCCTGGACCCCAATGTATACATCACGGG 692  
 Db 253 TGGCTCAAGAAATGAGATGTATCGACCCCAACCCAGGACACCAACTTCTCTGCTCACCATC 312  
 QY 693 GAGCAGCCTGGTGGGACAGCGCCCTTGTGACACGGCCAACTACACCTTGGCTG 752  
 Db 313 GACCAAACTCATCATCCGCGAGGCCCTGTCTGGACACTGCGCAACTATACCTTGGCTG 372  
 QY 753 GCCAAGAACTCTGTGGCACGCTCCCGTCCCGTCCCGTGTCTCATCTGTACGTGAAC 812  
 Db 373 GCCAAGAACTCTGTGGCCAAACCGCGAGACCACTGCCACCTCATCTGTACGTGAAT 432  
 QY 813 GGTGGTGTGCACTGAGACCGAGTGTCCGTCTGACGCGCAGCTGTGGGCGCGCTGG 872  
 Db 433 GCGGCTGTGTCAGCTGGCGAGAGTGTGTCACCTCTCCAAACCGCTGTGGCGAGGCTGG 492  
 QY 873 CAGAAACGAGCGCGAGCTGCACCAACCGGGGCTCTCAAGGGGGGCGCTTCTGTGAG 932  
 Db 493 CAGAAACGCGCGAGCTGCACCAACCGCGCTCCACTCAACGGAGGGGCGCTTCTGCGAG 552  
 QY 933 GGGCAGAACTGTCCATGACCGCACCGTCTCTCTGTCTGTCTGTGTCAGCGCAGCTGG 992  
 Db 553 GGCAGGCAATTCAGAA---GACGCGCTGCACCACTGTGCCAGTCAATGGGCGCTGG 609  
 QY 993 AGCCGTGGAGCAAGTGTGTGGCTGTGGCTGTGACACTGCACCCACTGCGGAGCGCTGAG 1052  
 Db 610 ACGGAGTGGAGCAAGTGTGTGAGCTGAGTGTGCTGAGTGTGCTGAGTGTGAGCGAG 669  
 QY 1053 TGCTCTGACCCAGCACCCCGCAACGGAGGAGAGTGCAGGCGACTGACCTGACACCC 1112  
 Db 670 TGCATGGCGCCCCACCCCAAGACGGAGGCGCTGACTGACGCGGCGCTGCTGACTCT 729  
 QY 1113 CGCAACTGTACAGTGAACCTCTGTGTACACAGTCTTCT----- 1151  
 Db 730 AAGAACTGCAGATGGCTGTGATGCATAATGAAGAACTTAAGGACCCCAACAGC 789  
 QY 1152 -----GGCCCTGAGAGCGTGGCCCTCTATGTGGGCGCTCT---ATCGCGTGGCC 1196  
 Db 790 CACCTGTGGAGGCTCAGGGATGCGGCGCTGTATGCGGGCTGCTGGTGGCTATCTTC 849  
 QY 1197 GTCTGCTGTGCTGTGCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1256  
 Db 850 GTGTCGTGGCAATCTCTCATGGGCTGGGGTGGTGTGTATACCGCCCAACTGCGCTGAC 909  
 QY 1257 CTGACTCAGATGTGGTGTGACTCGTCCATT---CTCACCTCAGCTTCCAGCCCGCTGAGC 1313  
 Db 910 TTGCACACAGACATCACTGACTCATCTGTGCGCTGACTGGTGTGTTCACCCCGTCAAC 969  
 QY 1314 ATCAAGCCCGAGCAAGAGCAGCAACCCCGATCTGTCT-----CACATCCAGCGGAGCTC 1367  
 Db 970 TTTAAGACGCGAGGCGCCAGCAACCCCGAGCTCTCTACACCCCTCTGTGCTCTCTGACTG 1029  
 QY 1368 AGCACACACACCTTACAGGGGAGTCTCTGTCTCCCGGAGGATGGGCGCCAGCCCCAAG 1427

Db 1030 ACAGCAGCGCGGATCTACCGGAGACCCGTGTATGCCCCTGCAGGACTCCACGACAAA 1089  
Qy 1428 TTCCAGCTCACAATGGCACTGTCTAGCCCC----- 1460  
Db 1090 ATCCCATGACCAACTCTCTCTGTCTGTGACCCCTTACCAGCCCTTAAGGTCAAGGTCTAC 1149  
Qy 1461 -----CTGGGTGGC 1469  
Db 1150 AGCTCCAGCACACCGGCTCTGGCCAGGCTTGGCAGATGGGGCTGACCTGTCTGGGGTC 1209  
Qy 1470 GCGCGCACACACTGACACAGCTCTCCACCTCTGAGGCGGAGGAGTGTCTCTCCGC 1529  
Db 1210 TTGCGCTGGACATACCTAGCAATTTGCGCGGAGACCCACTTCTGTGACCTGGC 1269  
Qy 1530 CTCTCCACCCAGAACTAC-----TTCCGCTCCCTGCCCGGAGGACACAGCAACATGACC 1583  
Db 1270 AGCGCAGCTCGGTTCCAGCAGCTCTTGGGCTTGGCCGAGACCCAGGAGCAGCTC 1329  
Qy 1584 TATGGGACCTTCAACTTCTCTGGGGCGGCTGATGATCCCTTAATACAGGTATGACCTC 1643  
Db 1330 AGCGCAGCTTTGGCTGTCTGGTGGGAGGCTCAGCATCCCGGACACAGGGGTGAGCTTG 1389  
Qy 1644 CTATCCCCCGAGATGCCATACCCCGAGGGAAGATCTATGAGATCTACCTCAGCTGCAC 1703  
Db 1390 CTGTGCCCCAATGGAGCCATCCCGAGGCAAGTTCTACGAGATGATCTACTCATCAAC 1449  
Qy 1704 AAGCGGAGAGCTGAGGTGTCCTTGGCTGTCTGAGCTGTCTGAGCCCTGCTGAGTCCCATGCTT 1763  
Db 1450 AAGCGAGAAAGTACCCTCCGCTTTTCAAGAGGAGCCAGACAGATTTAGAGCCCTCGGTG 1509  
Qy 1764 AGCTGTGACCCCTGGGCTGTCTGTCTACCCGGGACGATCATCTGGCTATGACCACTGT 1823  
Db 1510 ACCTGTGACCCACAGGCTCTGTCTGTGCGCCCGCTCATCTCACCATGCCCCACTGT 1569  
Qy 1824 GGGGAGCCAGCCCTGACAGCTGGAGCTGCGCTTCAAAAAGCAGTCGTCGAGGCGAC 1883  
Db 1570 GCCAAGTCAGTGCCTGCTGAGTCTTTAGCTCAAGCCAGGCCCCACAGGCGCAC 1629  
Qy 1884 TGGGAGCAGGATGTGTGACCTGGCGAGGAGGCGCCCTCCACCTCTACTACTGCGCAG 1943  
Db 1630 TGGG---AGGAGTGTGTGACCTGGATGAGGAGACCTGAACACACACCTGTACTGCCAG 1686  
Qy 1944 CTGGAGGCGAGCTGCTGCTGCTCTTACCGAGCAGCTGGCGCTTTCCTCTGGTGGGA 2003  
Db 1687 CTGGAGCCAGGCGCTGTACATCTCTGTGTCGACGCTGGGACCTACGTGTCTACGGGC 1746  
Qy 2004 GAGGCTCAGGCTGCTGCGCCAGCGCCTCAAGCTGCTCTGTGCGCGGTGGCC 2063  
Db 1747 GAGTCTATTCGCTCAGCAGTCAAGCGCTCCAGCTGGCGCTTTCGCCCCGCGCTC 1806  
Qy 2064 TGCACCTCCCTCAGTACAACTACCGGCTTACTGCTGTGATGACACCCAGCATGACTC 2123  
Db 1807 TGCACCTCCCTCAGTACAGCTCCGCTTACTGCTGGAGGACACGCTGTAGCACTG 1866  
Qy 2124 AAGGAGTGTGAGCTGAGAGCAGCTGGGGGACAGCTGATCAGGAGCCACGGCTC 2183  
Db 1867 AAGGAGTGTGAGCTGAGGAGCTCTGGCGGATCTTGTGAGGAGCGGAAACCG 1926  
Qy 2184 CTGCACCTCAAGACAGTGTACCAACCTCGCGCTATCCATCCAGCATGTGCCAGCTCC 2243  
Db 1927 CTAAATTTCAAGACAGTATACCAACCTCGCGCTCTCCCTCATGACCTCCCGCATGCC 1986  
Qy 2244 CTGTGAAAGTAACTCTCTGTGCTACGCTACGAGGATCCCTTTTATCATCTGGAAT 2303  
Db 1987 CATTTGAGGAGCAAGCTGTGGCGAAATACAGGAGATCCCTTCTATCATCTTTGGAGT 2046  
Qy 2304 GGCACGCGGTACTTGTGACTGACCTTCACTGAGCTGTGAGCTGTGAGCCCGACAGT 2363  
Db 2047 GGCAGCCAGAGGCCCTTCACTGACTTTTCACTGAGGAGGACAGCTTGGCTCCACA 2106  
Qy 2364 GACCTGGCTTGCAGCTGTGGGTGTGGCAGGTGGAGGCGACCGGAGAGCTTTCAGATC 2423

Db 2107 GAGCTCACCTGCAAGATCTGGTGGCGCAAGTGGGAAGGGGGCCAGATATTCAGCTG 2166  
Qy 2424 AACTTCAACATCAACAAG---GACACAGGTTTGTGAGCTGCTGGCTCTGGAGAGTGAA 2480  
Db 2167 CATACCACTCTGGCAGAGACACCTGCTGGCTCCCTGGACACTCTCTGTCTGCCCCCTGGC 2226  
Qy 2481 GCGGGGTCCAGAGCCCTGGTGGGCCCCAGTGCCTTCAAGATCCCCCTTCTCATTTGGCAG 2540  
Db 2227 AGCACTGTCAACCCAGCTGGAGCTTATGCTTCAAGATCCCACTGTCATCCGCCAG 2286  
Qy 2541 AAGATAATTTCCAGCTGGACCCACCTGTAGGCGGGTGGCGACTGGCGGACTCTGGCC 2600  
Db 2287 AAGATATCAACAGCTAGATGCCCCCACTCAACGCGGCAATGACTGGCGATGTAGCA 2346  
Qy 2601 CAGAACTCCAGCTGGACAGCATCTCAGCTTCTTTGCTCCAGCCAGCCCCCAGCC 2660  
Db 2347 CAGAAGCTCTTATGAGCCGCTACCTGAATTTTGGCCACCAAGCAGGCCCGCTGGT 2406  
Qy 2661 ATGATCTTCAACCTGTGGAGCGCGGCTTCCCAACGCGCACTCAGCCAGCTGGCT 2720  
Db 2407 GTGATCTTGGACCTCTGGAGAGCTCTGAGCAGGAGCTGGGAGCTCAACAGCTGGCG 2466  
Qy 2721 GCAGCAGTGGTGGAGCTGGCGCAGCCAGCGCTGGCTCTTTCACAGTGTGGAGGCTGAG 2780  
Db 2467 AGTGCCTTGGAGGAGATGGCGAAGAGTGAATGCTGTGGTGTGGCCACCGACGGGAC 2526  
Qy 2781 TGCTGA 2786  
Db 2527 TGCTGA 2532

## RESULT 6

AY406492

## LOCUS

Pan troglodytes UNCS gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

ACCESSION

AY406492

VERSION

AY406492.1 GI:39762466

KEYWORDS

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

Pan troglodytes

REFERENCE

1 (bases 1 to 2802)

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,  
Adams, M.D. and Cargill, M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios

TITLE

Science 302 (5652), 1960-1963 (2003)

JOURNAL

14671302

PUBMED

2 (bases 1 to 2802)

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,  
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,  
Perriera, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J.,  
Adams, M.D. and Cargill, M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios

AUTHORS

Direct Submission

Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering  
them based on alignment.

FEATURES

Location/Qualifiers

source

1..2802

/organism="Pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

&lt;1..&gt;2802

/gene="UNC5C"

/locus\_tag="HCM2575"

ORIGIN

Query Match

26.5%; Score 763.2; DB 29; Length 2802;

Best Local Similarity 52.9%; Pred. No. 8e-118; Matches 1456; Conservative 0; Mismatches 1218; Indels 78; Gaps 7;			
QY	98	GCOCGCCCTGTGGCCAGCGCTCTCGGGATAGTCTCGCGGTGGCTCGCGGCTCGGG	157
Db	66	GCTCGTGCTACTCGCGCTGGCCCTGCTCAGCGCCAGCGCACTGGCTCCGCCGCCAAGA	125
QY	158	TGCCACGACAGTGCACCGTGCCCAACCCAGTGCCTGTGCAACCCCGACCTCTCTCC	217
Db	126	TGATGACTTTTTCATGAATCCCAAGAACTTTTCTTCGGATCCACTGAGCCTCTGCC	185
QY	218	CCACTTCTCTGTGGAGCCCGAGGATGTGTACATCGTCAAGAAACCAAGCCAGTGTCTGT	277
Db	186	ACATTTCTTATTGAACCTGAGGAAGCTTATTATTGTGAAGAAATAGCCTGTCAACCTGTA	245
QY	278	GTGCAAGGCC-----GTGCCGCCACGACGATCTTCTTCAAGTGCAACGGGGAGTGGT	331
Db	246	CTGTAAAGCAAGCCCTGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	305
QY	332	GCGCCAGTGGACCACTGTATCGAGCGCAGCACAGACGGGAGCAGTGTGAGCGCACCAT	391
Db	306	NN	365
QY	392	GGAGTCCGCAATTAATGTCTCAAGCAGCAGTTCGAGAAGGTGTTCGGGCTGGAGGAATA	451
Db	366	NN	425
QY	452	CTGGTGCCAGTCCGTGGCATGGAGCTCTCTCGGGCACCAAGAGTCAAGAGCCCTACAT	511
Db	426	NN	485
QY	512	CCGCATAGCCAGATTGCGCAAGAACTTCGACAGAGCGCTGGCCAGGAGGTGTCCT	571
Db	486	NN	545
QY	572	GGAGCAGGCATCGTGTGCTGCTGCGCTCCACCGGAGGGATCCCTCCAGCCGAGTGA	631
Db	546	GGAACAGGAAGTCTTACTTCAGTGTGGCCACTGAAAGGATCCAGTGGCTGAGNNNN	605
QY	632	GTGGCTCCGGAAACGAGGACCTGTGGACCGCTCCCTGGACCCCAATGTATACATCAAGCG	691
Db	606	NN	665
QY	692	GGACACAGCCTGTGTGTGCGACAGGCCCGCTTGTGACAGCGCAACTACACCTGCGT	751
Db	666	TGATCACAACTCATATAAGCAGGCGCCGACTCTCTGATCTGCAATTTATACCTGTGT	725
QY	752	GGCCAAAGAACATCGTGGCAGCTGCCCGCAGCGCTCCGCTGCTGCTCATCGTCTACGTGA	811
Db	726	TGCCAAAAACATTTGTGCCAAGAGGAAAGTACAACTGCCACTGTCTATGCTATGTCAA	785
QY	812	CGGTGGTGTGCAAGTGGACCGAGTGGTCCGTCTGACGCGCCAGCTGTGGCGCGCGCTG	871
Db	786	CGGTGGCTGTGCTCACCTGGAACGAGTGGTCTGTGTGTAAACAGCGCTGTGGAACGAGGA	845
QY	872	GCAGAAACGGACCGGAGCTGCACCAACCGCGCTCTCAACGGGGCGCTTCTGTGA	931
Db	846	TCAGAAACGTACAAGGACTTGTACCAACCGGACCACTCAATGGGGGTGCTCTGTGA	905
QY	932	GGGCGAAGATGTCATGACCGCACCGCTCTCTCTGCTGTGCTGTGAGCGGAGCTG	991
Db	906	AGGCGAGTGTGCAGAA---AATAGCCTGTACTACGTTATGCCAGTGGATGGCAGGTG	962
QY	992	GAGCCGTGGAGCAAGTGTGGCTGTGGCTGGACTGCACCCACTGGCGGAGCCGTGA	1051
Db	963	GACGCCATGGAGCAAGTGTCTACTTGTGAACTGAGTGCACCCACTGGCGCAGGAGGA	1022
QY	1052	GTGCTCTGACCCAGCACCCCGCAACGGGAGGAGTGCACCGGCACTGACCTGGACAC	1111
Db	1023	GTGCAACCGCCAGCCCCCAAGAAATGGAGCAAGGACTGCGATGGCTCTGTGCAATC	1082
QY	1112	CCGCAACTGTACAGTGAACCTCTGTGTACAGTGTCTTGGCCCTGAGGACGTGGCCCT	1171
Db			
1083	CAAGAACTGCACGTATGGGCTTTGCATGTCAGACTGCTCTCTGATTCAGATGATGTGCTCT	1142	
QY	1172	CTATGTGGG---CCTCATPGCGGTGGCGCTCTGCTGCTGCTGCTGCTGCTGCTCTCAT	1228
Db	1143	CTATGTTGGGATTTGTATAGCAGTGTGTTTGGCTGGCGATCTCTGTAGTGTGGCCTT	1202
QY	1229	CCTCGTTTATTTGCCGAAGAGGAGGCTGGACTCAGATGTGGCTGACTGCTGCTCATCTCT	1288
Db	1203	GTTTGTGTATCGGAAGAAATCATCGTGAATTTGAGTCAGATATTTATGACTCTTCGGCACT	1262
QY	1289	CACCTCAGGCTTCCAGCCCGTCAAGCTCAAGCCAGCAAGCAGCAGACACACCCCATCTGCT	1348
Db	1263	CAATGGGGCTTTTTCAGCCTGTGAACATCAAG-----GCAGCAAGCAACAGATCTGCT	1313
QY	1349	CACCATCCAGCCGAGCCTCAGCACCAACCACTTACAGGGCAGTCTCTGTCCCCGGCA	1408
Db	1314	GGCTGTATACCCCGAGACCTCAGCTCAGCTCAGCCATGTACAGAGGACCTGTCTATGCCCT	1373
QY	1409	GGATGGCCCGAGCCCAAGTTCCAGCTCAGCAATGGGCACTGCTCAGCCCGCTGGGTGG	1468
Db	1374	GCATGACGTCTCAGACAAATCCCAATGACCAACTCTCCAATTTCTGGATCCACTGCCCAA	1433
QY	1469	CGCGCCGCACACACTGCACACAGCT-----CTCCCACTCTCAGGCCGAGGA	1516
Db	1434	CCTGAAATCAAAAGTGTACAACCTCAGGTGCTGTACCCCCCAAGATGACCTCTCTGA	1493
QY	1517	GTTCTGCTCCGCTCTCCACCCAGAACTACTTCCGCTCCCTGCCCGAGGACCCAGCAA	1576
Db	1494	GTTTACGTCCAAGTGTCCCTCAGATGACCCAGTCATTGTTGGAGATGAAGCCTCAG	1553
QY	1577	CATGA-----CTATGGGACCTT	1594
Db	1554	CCTGAAGAACACAGTCTAGCAAGGACAGCTGATCCATCTGTACCGCATTTGGCAGCTT	1613
QY	1595	CAACTTCTCGGGGCGGCTGATGATCCCTAATACAGGTATCAGCCTCTCTCATCCCCC	1654
Db	1614	CNNNTGCTGGNNNNACCTTATTTGTTCCCAATTCAGAGTCAGCTGCTGATTTCCCGC	1673
QY	1655	AGATGCCATACCCGAGGAAAGATCTATGAGTCTTACCTTACCTGTCACAGCCGGAAGA	1714
Db	1674	TGGGCCCATTTCCCAAGGAGAGTCTACGAATGTATGTACTGTACACAGGAAAGAAAC	1733
QY	1715	CGTGAAGTGGCCCTAGCTGGCTGTGACACCTGCTGAGTCCCATCTGTTAGCTGGACC	1774
Db	1734	TATGAGGCCACCCATGGATGACTCTCAGACACTTTTGAACCTGTGTGAGTGTGGGCC	1793
QY	1775	CCCTGGCGTCTGCTCACCCGCGCAGTCTCTCTGGCTATGACCACTGTGGGAGCCGAG	1834
Db	1794	CCGAGGAGTCTGTCTCACCCGCGCGCTGCTCTTACTATGTCATCTGCGCAGACCCCAA	1853
QY	1835	CCCTGACAGCTGGAGCCTGCGCTCAAAAAGCAGTCTGTCGAGGGCAGCTGGAGCAGGA	1894
Db	1854	TACCGAGGACTGGAAAAATCTGCTCAAGAACACAGGACAGCAGGACAGTGGG---AGA	1910
QY	1895	TGTCTGCACTGGCGGAGGAGCGCCCTCCCACTTACTTACTTGCAGCTGGAGGCGCAG	1954
Db	1911	TGTGGTGGTGGTTCGCGGAGGAAAACTTCAACACCCCTGCTTACATTCAGCTGGATGAGA	1970
QY	1955	TGCTGTCTACGTCTTTCACCGAGCAGTGGCGCTTTGGCTGTGGGAGGCGCCTCAG	2014
Db	1971	GGCTTGCCACATCTCTCAGAGAACCTCAGCACCTTACGCCCTGTGAGCAATTCACCAAC	2030
QY	2015	CGTGTCTGCCCAAGCGCTCAAGCTGCTTCTGTTTGGCGCGGTGGCTGCACTCCCT	2074
Db	2031	CAAAAGCGCTGGAAAGCGCTCAAGCTGGCCATCTTTTGGGCCCTTGTCTGCTCTGCT	2090
QY	2075	CGAGTCAACATCCGGGTCTACTGCTCTGATGACACCCAGATGCACTCAAGGAGGTGGT	2134
Db	2091	GGAGTACAGCATCCGAGTCTACTGTCTGGATGACACCCAGGATGCCCTGAAGGAAATTT	2150
QY	2135	GCAGCTGGAGAGCAGCTGGGGGACAGTGTATCCAGGAGCCACGGTCTCTGCACTTCAA	2194
Db	2151	ACATCTTGAGAGACAGATGGGAGGACAGCTCTTAGAAGAACCTTAAGGCTCTTCAATTTAA	2210

QY	2195	GGACAGTTACCAAACTGCGCCTATCCATCCACGATGTGCCAGCTCCCTGTGAAGAG	2254
Db	2211	AGGACAGCCCAAACTGGCCTGTCAATTACGATATCGCCCATTCCTCTGGAGAG	2270
QY	2255	TAAGCTCCTTGTGTCAGCTACACGAGATGCCCTTTTATCATCTGGAATGGCACGACGG	2314
Db	2271	CAAAATTCCTGGCTAAATATCAGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	2330
QY	2315	GTACTTGCACTGCACTTTCACCTTGGAGCTGTCAGCCACGACACTAGTACCTGGCTG	2374
Db	2331	NAACCTGCACTGCACTTTCATCTCTGGAAGATTAGCCTGAACACATGGAGCTGGTTG	2390
QY	2375	CAGACTGTGGTGTGGCAGCTGGAGGCGACGGCGAGAGCTTCAGCATCAACTTCAACAT	2434
Db	2391	CAAACTCTGTGTGCGGCAGCTGGGAAGGAGGCGAGATCTTCAGCTCAACTTGCACGT	2450
QY	2435	CACCAAGCACAAAGGTTTGTGTGAGCTGCTGGGCTCTGGAGAGTGAAGCGGGGTCGCCAGC	2494
Db	2451	GTACAGGAACCTACTGGCATCGATTTCGCTCTGGATCCTCGGAACACCATCACCAC	2510
QY	2495	CCTGGTGGGGCCAGTCCGTTCAAGATCCCTTCCTCATTTCTGGCAGAAAGATATTCCAG	2554
Db	2511	GGTCACGGGGCCAGTGTCTTCAGCATCCCTCTCCCTATCCGGCAGAAAGCTCTGTAGCAG	2570
QY	2555	CCTGGACCCACCTGTGTAGCGGGGTGGGACTGGCGGACTCTGGCCCGAAACTCCACCT	2614
Db	2571	CCTGGATGCCCCCAGACGAGAGGCCATGACTGGAGGATGCTGGCCCATGAAGNNNNNN	2630
QY	2615	GGACAGCATCTCAGCTCTTTTGCTCCAAAGCCAGCCCAAGCCCATGATCTCAACCT	2674
Db	2631	NNNNNGTACTTGAATTACTTTTGCCACCAATCCAGGCCAACTGGCGTAACTCTGGATCT	2690
QY	2675	GTGGAGCGCGGCACCTTCCCAACGGCAACCTCAGCCAGCTGGCTGCAGAGTGGCTGG	2734
Db	2691	TTGGGAAGCACAAGAACTTCCAGATGGAAACCTTCAGCATGCTGCGACGTGTCTTGGAGA	2750
QY	2735	ACTGGGCGACGACGCTGGCCCTTTCAGAGTGTGGAGGCTGAGTGTGCA	2786
Db	2751	AATGGGAAGACATGAACCGGTGGTGTCTCTAGCAGCAGAGGGCAGATTTAA	2802

## ORIGIN

Query Match	26.2%	Score 756.2	DB 12	Length 1034	
Best Local Similarity	89.7%	Pred. No. 7.5e-117			
Matches 905	Conservative	0	Mismatches 88	Indels 16	Gaps 8
191	GCCTGTCGCCAACCGGACCTGTTCCCCACATCTCTCGTGGAGCCGAGGATGTACAT	250			
1	GCCTGGTGCACACCGGACCTGTTCCCACTTCTCTGGTGGAGCCGAGATGTACAT	60			
251	CGTCAAGAACAAAGCCAGTGCTGCTTGTGTGCAAGCCGCTGCCGCCACGAGATCTTCTT	310			
61	CGTCAAGAACAAAGCCAGTGCTGCTTGTGTGCAAGCCGCTGCCGCCACGAGATCTTCTT	120			
311	CAAGTGCACGGGGAGTGGTCCGCCAGGTGCACACAGCTGATCGAGCGACGACACGCG	370			
121	CAAGTGCACGGGGAGTGGTCCGCCAGGTGCACACAGCTGATCGAGCGACGACACGCG	180			
371	GAGCAGTGGTGAAGCCGACCATCGAGGTCCGCGATTAATGTCTCAAGGCAGCAGGTCTGAGAA	430			
181	GAGCAGTGGGCTGCCACCATCGAGGTCCGCATTAATGTCTCAAGGCAGCAGGTCTGAGAA	240			
431	GGTGTTCGGGCTGGAGGAATCTGGTCCAGTCCGTGGCATGAGCTCTTCGGGCACAC	490			
241	GGTGTTCGGGCTGGAGGAATCTGGTCCAGTCCGTGGCATGAGCTCTTCGGGCACAC	300			
491	CAAGAGTCAGAGGCCCTACATCCGCATAGCCAGATTGCCGAAGAACTTCGAGCAGAGGCC	550			
301	CAAGAGTCAGAGGCCCTACATCCGCATAGCCATTTTCGCGAAGAACTTCGAGCAGAGGCC	360			
551	GCTGCGCAAGGAGGTGTCCTCGAGCAGGCGCATCGTCTGCCCTGCCGTCACCGCGAGGG	610			
361	GCTGCGCAAGGAGGTGTCCTCGAGCAGGCGCATCGTCTGCCCTGCCGTCACCGCGAGGG	420			
611	CATCCCTCCAGCCGAGGTGGAGTGGTTCGCGAACGAGGACCTGGTGAACCGCTCCCTGGA	670			
421	CATCCCTCCAGCCGAGGTGGAGTGGTTCGCGAACGAGGACCTGGTGAACCGCTCCCTGGA	480			
671	CCCCAATGTATACATCACCGGGAGCACACCTGTTGGTGGACAGGCCCGCCTTGTGTA	730			
481	CCCCAATGTATACATCACCGGGAGCACACCTGTTGGTGGACAGGCCCGCCTTGTGTA	540			
731	CACGGCCAACTACACCTGCGTGGCCCAAGAACATCGTGGCAGCTCGCCGACAGCGCTCCGC	790			
541	CACGGCAACTACACCTGCGTGGCCCAAGAACATCGTGGCAGCTCGCCGACAGCGCTCCGC	600			
791	TGCTGTTCATCGTCTACGTGAACCGTGGTGGTTCGA-CGTGGACCGAGTGGTTCGTCTGCA	849			
601	TGCTGTTCATCGTCTACGTGAACCGTGGTGGTTCGA-CGTGGACCGAGTGGTTCGTCTGC-	659			
850	GGCCAGCTGTGGGCGCGCTGCGAGAAACCGGAGCCGAGCTGCACACCCGGTGCCTC	909			
660	GGCCAGCTGTGGGCGCGCTGCGAGAAACCGGAGCCGAGCTGCACACCCGGTGCCTC	719			
910	TCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCCAATGACCGCACCGCTCTCTCTGTCG	969			
720	TCAACGGGGCGCTTTCTGTGAGGGGAGAAATGTCCAATGACCGCACCGCTCTCTCTGTCG	774			
970	TTGCTCTGTGACCGCAGCTGGAGCCGCTGGACCAAGTGGTTCGCGCTGT--GGCTCGGA	1021			

```

Db      775 CTGTGCCAGTGGACGACGCTGTAGCACGTGGAGCCAGTGGTGGGCTTGGC 834
QY      1028 CTGACCCACT--GGCGGACCGGTGAGTGTCTGACCCAGCACCCGCGACGAGGGAG 1085
Db      835 TTGACCCACTTGGCGGAGCCGGAGTGTCTGAACCCAGCACCCCGGACGCGAGGGG 894
QY      1086 GAGTG---CCAGGCACTGACCTGGACACCCGCAACTGTACCAGTGAACCTCTGTGTAC 1142
Db      895 AGTGTGCGCCAGTCACTGGACCTGGACCCCGGA-TGTTCCAGTGAAGTCTGTGT-CCC 952
QY      1143 AGTGTCTTCTGGCCCTGAGACGCTGGCCCTCTATGTGGGCTCATGCCG 1191
Db      953 ACGGGTTCTGGCCCTGAGGACTTGGCTCTCTATGTGGGCTCATGCCG 1001

RESULT 8
LOCUS   AY411749
DEFINITION Mus musculus HCM4327 gene, VIRTUAL TRANSCRIPT, partial sequence,
AY411749
ACCESSION AY411749
VERSION
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)

REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source 1..2532
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>2532
/locus_tag="HCM4327"

ORIGIN
Query Match 25.2%; Score 726.4; DB 29; Length 2532;
Best Local Similarity 59.1%; Pred. No. 1.2e-111;
Matches 1492; Conservative 0; Mismatches 896; Indels 138; Gaps 9;

QY      393 GAGTCCGCATTATGTCTCAGGACGAGTCGAGAGGTGTTCCGGCTGGAGGAATAC 452
Db      13 GAGTGCAGATCGAGGTGTACGGCAGCAAGTGGAGGAATCTTCCGGCTCGAGGACTAC 72
QY      453 TGGTGCCAGTCCGTGGCATGAGTCTCTCGGACCCAGGAGTCTCAGAGGCTTACATC 512
Db      73 TGGTGCCAGTCCGTGGCTGAGTCTTTCGGAACTACCAAGAGTCCCGGACCTACATC 132
QY      513 CGCATGCCAGATTGCGCAAGAACTTCGAGCAGGAGCGCTGGCCACAGGAGGTGTCCCTG 572
Db      133 CGCATTCGCTACTTGGCGAAGAACTTTGACAGAGGCTCTGGCCAGAGGAGTACCTTTG 192
QY      573 GAGCAGGGCATGTGTGCTCCCTGCGTCCACCGGAGGCGATCCCTCCAGCGGAGGTGGAG 632

```

```

Db      193 GATCATGAGGTCTTCTGTGAGTCCGCCCCACCGAGGGAGTGCCTGTGGCTGAGGTGGAA 252
QY      633 TGGCTCCGGAACGAGGACCTGGTGGACCCGTCCCTGGACCCCAATGTATATCATCATCCGCG 692
Db      253 TGGCTCAAGAAATGAAGATGTCTATTGACCCCGCTCAGACACATACTTCTGTCTCACCATT 312
QY      693 GAGCAGAGCTGGTGGTGGAGAGGCGCCCTTGTGTGACGCGGCAACTATACCTCGGTG 752
Db      313 GACCAACAATCATCATCCGCGAGGCGCGCTCTCAGACACGCGCAACTATACCTGTGTG 372
QY      753 GCCAAGAAATCGTGGCACGTCGCGCAGCGCTCCGCTGTCTGTGTCATCTGTACCTGTAAC 812
Db      373 GCCAAGATATCGTGGCAAGCGCGGAGCACACGCGCCACAGTCACTGTCTATGTGAAT 432
QY      813 GGTGGTGTGTGACGCTGGACCGAGTGGTCCGTCTGTGACGCGCAGCTGTGGGCGCGCTGG 872
Db      433 GGAGGCTGTGTCAGCTGGGCGAGTGGTCAACCTGTTCCTCAATCGCTGTGGCGAGGCTGG 492
QY      873 CAGAAACGAGCCGAGCTGTGACCAACCGCGGCTCTTCAACGCGGCGCTTCTGTGTGAG 932
Db      493 CAGAAGCGTACTCGGACCTGCAACCAATCCAGCCCCACTCAATGAGAGCGCTTCTGTGAG 552
QY      933 GGGCAGAAATGTCCATGACCGCACCGCTCTCTCTGTGTGTGTCTGTGACGCGAGCTGG 992
Db      553 GGACAGGCTTCCAGAA---GACAGCTTGACACCGGTGTGCCAGTGGATGGAGCGTGG 609
QY      993 AGCCGCTGGAGCAAGTGTGGGCTGTGGGTGGACTGCACCCACTTGGCGGAGCGGTGAG 1052
Db      610 ACCGAGTGGAGCAAGTGTGTCTGCTGTGACGACAGAGTGTGCGCACTGGCGCAGCCGAG 669
QY      1053 TGCTCTGACCCAGCACCCGCAACGAGGAGGAGTGGCAGGCACTGACCTGACAC 1112
Db      670 TGATGGCAGCCGCAACCCGAGAGGCGCGTGTGATGACGCGGAGCGCTACTTCACTCC 729
QY      1113 CGCACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCTTGAGGAC--- 1163
Db      730 AAGAACTGCACTGATGGCTGTGCTGTGTGATCAGAGAACTCTTAAAGCACTTAAAGC 789
QY      1164 -----GTGGCCCTCTATGTGGGCTCATCGCGCTGGCCCTG 1199
Db      790 CACCCCTGGAGACATCGGAGATGTGCACTGTACGCGAGGCTTGTGTGGCTCTTT 849
QY      1200 TGCTGTGCTCTGCTGCTGTCTCTCT---CATCTCTGTTTATGCGGAGAGAGAGGG 1256
Db      850 GTGGTGGTGGGTTCTCATGGCGTGGAGTGTATATACCGGAGAACTTCCCGGAC 909
QY      1257 CTGACTCAGATGTGGTGAACCTCT---CCATTTCTCACTCAGGCTTCCAGCGCTGACG 1313
Db      910 TTCGACACGAGACATCACCGACTCTCTGCGGCTCTACTGTGTGCTTCCACCTGTCAAC 969
QY      1314 ATCAAGCCCGCAAGAGCAGCAACCCCATCTGTCT-----CACCATCCAGCCGACCTC 1367
Db      970 TTCAAGACTGCAAGGCCCAACCAACCCGAGCTCTCTGCACCGCTCCGCCCTCCAGACTA 1029
QY      1368 AGCACCACCACTTACAGGAGGAGTCTCTGTCCCGGAGGATGGGCCAGGCCCAAG 1427
Db      1030 ACGGCCAGTGTGTCATCTACCGCGGCTCTGTGTATGCCCTGCGAGGACTTCCCGGACAAG 1089
QY      1428 TTCCAGCTCACCAATGGGCACCTGTCT----- 1453
Db      1090 ATCCCATGACTAATTCGCCCTCTGTGGATCCCTGCGGAGCTCAAGATCAAGGTCTAT 1149
QY      1454 -----CAGCCCTCTGGGTGGC 1469
Db      1150 AACTCCAGCACCATCGGTCTCTGGCTCTGGCTGGTGTGATGAGCCGAGCTGTGGGTGTC 1209
QY      1470 GCGCGGCACACACTGCACACAGCTCTCCACCTCTGAGCGGAGGAGTGTGCTTCCCGC 1529
Db      1210 CTCGCGCGGCGACGTACCCAGGCGATTTCTCCGGGACACCCATTTCTGTGACCTCGC 1269
QY      1530 CTCTCCACCCAGAACT-----ACTTCCGCTCCCTGCCCGGAGCACCAGCAATGACC 1583
Db      1270 AGTGCCAGCCTTGGTTCCCGAGCACTCTCTGGGCTTACCTCGGAGCCCGACGACGAGTGC 1329

```

QY 1584 TATGGACCTTTAACTTCCTCGGGGCGGTGATGATCCCTAATACAGGATATCAGCCTC 1643  
Db 1330 AGCGGACCTTTGGTTGGCTGGAGGAAGCTGAGCTTCCCGGCACAGGGGTGAGCCTG 1389  
QY 1644 CTATATCCCCCAGATGCCATACCCCGAGGGAAGATCTATAGATCTACCTCAGCGTGCAC 1703  
Db 1390 TTGTTACCAATGGAGCCATTTCCCGAGGCAAGTTCTATGACCTGTATCTACATATCAAC 1449  
QY 1704 AAGCGGAAGACGTGAGGTGGCTTGGCTGTGAGACCCCTGCTGAGTCCCATCGTT 1763  
Db 1450 AAGGCGGAAGACACCTCCCATTTTCAAGAGTTCCAGACAGTATTGAGCCCTCGGTG 1509  
QY 1764 AGCTGTGAGACCCCTGGCGTCTGTCTACCCGCGCAGTCACTCTGGCTATGAGACCACTGT 1823  
Db 1510 ACCTGTGGGCCCCACAGGCTACTCTCTGTGCGCGCTGTCTCTCACCGTGCACCACTGT 1569  
QY 1824 GGGAGCCAGCCCTGACAGCTGGAGCCTGCGCTCAAAAAGCAGTCTGTGGAGGGCAGC 1883  
Db 1570 GCTGAAGTATCGCTGGAGACTGGATCTTTAGCTCAAGACCCAGGCCCATCAGGGCCAC 1629  
QY 1884 TGGAGACAGGATGTCTGCACCTGGCGGAGGAGGCGCCCTCCACCTCTTACTTACTGCGAG 1943  
Db 1630 TGGG---AGGAGTGTGACCTTGGATGAGGAGACCCCTCAACACACCTGTCTACTGCGAG 1686  
QY 1944 CTGGAGGCGAGTGTCTGTCTTACCGAGAGCTGGGCGGCTTGGCCCTGTGGGA 2003  
Db 1687 CTGGAGGCTAAGTCTGCGCACATCTCTGTGGACCACTGGGTGACCTACGATTTCAATGGGC 1746  
QY 2004 GAGGCGCTCAGCGTGGCTGCGCCAGCGCTCAAGCTGCTTCTGTGCGCGGTGGCC 2063  
Db 1747 GAGTCTTACTCTCGTCTGCACTGAAGCGGCTCCAGCTGGCCATCTTGGCCAGCCCTC 1806  
QY 2064 TGCACTCTCCTCGAGTACACATCCGGGTCTACTGCTGCATGACACCCACGATGACATC 2123  
Db 1807 TGCACTCTCCTCGAGTATAGCTCAGGCTCTACTGCTGGAGGACACACTGTAGCACTG 1866  
QY 2124 AAGGAGGTGGTGAAGTGGAGAGCAGTGGGGGACAGCTGATCCAGAGGACACGGTTC 2183  
Db 1867 AAGGAGGTCTGAGCTGGAGAGGACTCTGGGTGGCTACTTGGTGGAGGAGGCCAAGCCT 1926  
QY 2184 CTGCACTTCAAGGACAGTTTACCAACCTCGCGCTATCCATCCAGCATGTGCCAGCTCC 2243  
Db 1927 TTGCTCTTTAAGACAGTTTACCAACCTAAGGCTCTCCCTCCATGACATCCCCCATGCC 1986  
QY 2244 CTGTGGAAGAGTAAGCTCTGTGAGCTACAGGAGATCCCTTTTATCAATCTGGAAT 2303  
Db 1987 CACTGGAGGAGCAAACTACTGGCCAAAGTACAGGAGATTCCCTTCTACCAAGCTCTGGAAT 2046  
QY 2304 GGCAGCAGGGTACTTGCACTGACCTTCACTGGAGGTGTGAGCCCGCAGCACTAGT 2363  
Db 2047 GGCAGCAGAGAGCCCTGCACTGCACTTTACCCCTGGAGGAGCATAGCTGGCCCTCAGC 2106  
QY 2364 GACCTGGCTGCAAGCTGTGGTGTGGCAGGTGGAGGCGACGGGAGAGCTTCAGCATC 2423  
Db 2107 GAGTTCACTTGAAGTCTGCTGCGGAGCTGCAAGGGAAGGAGGAGATTTCCAGCTG 2166  
QY 2424 AACTTCAACATCAC---AAGGACACAAAGTTTGTGAGCTGTGCTGTGAGTCTGAGAGTGAA 2480  
Db 2167 CACACACAGTTGGCCGAGACGCTGTGCTGCTCCCTGGATGCTCTGTCTGCCCCGGGC 2226  
QY 2481 GCGGGGTCCAGCCTGTGGTGGGCGCCAGTGCCTTCAAGATCCCTTCTCATTTCCGCGAG 2540  
Db 2227 AATGCCATCACCCAGCTGGGACCCCTATGCTTCAAGATACCCCTGTCTCATTCGCCCAA 2286  
QY 2541 AAGATATTTCCAGCTTGGACCCACCTGTAGGCGGGGTGCCGACTGGCGACTCTGGCC 2600  
Db 2287 AAGATCTGAGAGGCTTGGAGCGCCCCCACTCCGGGGCAACGACTGGAGGCTGTGGCG 2346  
QY 2601 CAGAACTCCACTGGGACAGCCATCTCAGCTTCTTGTGCTTCAAGGCCAGCCCAAGCC 2660  
Db 2347 CAGAGCTGTGCTGAGCGGCTACCTAACTACTTGGCCACCAAGAGTAGTCCCAAGCT 2406

QY 2661 ATGATCTCAACTGTGGAGGCGGCGCACTTCCCAACGGCAACCTCAGCCAGCTGGCT 2720  
Db 2407 GTATCTTAGACCTCTGGAGCTCGCAACAGGATGACGGGACCTCAACAGCCTGGCC 2466  
QY 2721 GCAGCAGTGGCTGACTGGGCGCAGCAGCCTGCTTTCACAGTCTCGAGGCTGAG 2780  
Db 2467 AGTGCTTGGAGGAGATGGGCAAGAGTGAGATGCTGTGGTAGCCAGTGGCCACAGATGGCAT 2526  
QY 2781 TGCTGA 2786  
Db 2527 TGCTGA 2532

RESULT 9  
BX364574 309 bp mRNA linear EST 05-MAY-2003  
LOCUS BX364574 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
DEFINITION CDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.  
ACCESSION BX364574  
VERSION BX364574.1 GI:30368812  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 909)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Genoscope - Centre National de Sequencage  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3529.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS1AC006ZF06Q1&cluster=3529.f. Contact :  
Peng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Paradise Avenue Genoscope sequence ID : CS1AC006ZF06Q1.

FEATURES  
Location/Qualifiers  
1..909  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC023YB10"  
/tissue\_type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 25.2%; Score 726.2; DB 13; Length 909;  
Best Local Similarity 98.4%; Pred. No. 7,9e-112;  
Matches 739; Conservative 4; Mismatches 7; Indels 1; Gaps 1;  
QY 1 AGTGGGGTCCGGGTGAGGGCTAAAGCGCCCTCCCGCGGGGGCGCCCGCCCG 60  
Db 160 AGTGGGGTCCGGGTGAGGGCTAAAGCGCCCTCCCGCGGGGGCGCCCGCCCG 219  
QY 61 GCCCGCCCGCTCCCGCGGGCCATGCGCTCGCGCCCGCGCTGTGCCAGCGCTCC 120  
Db 220 GCCCGCCCGCTCCCGCGGGCCATGCGCTCCCGCGGGGGCGCCCGCGCTCC 279  
QY 121 TGGGATAGTCTTCGCGGCTTGGCTCCCGGGTTCGGGTGCCAGAGTGCACCGTGG 180  
Db 280 TGGGATAGTCTTCGCGGCTTGGCTCCCGGGTTCGGGTGCCAGAGTGCACCGTGG 339  
QY 181 CCAACCAAGTGTGTGGCAACCCCGGACTGTCTCCCACTTCTCTGGTGGAGCCGAGG 240





RESULT 11  
 BC033727  
 LOCUS  
 DEFINITION Homo sapiens netrin receptor Unc5h1, mRNA (cDNA clone IMAGE15166762), containing frame-shift errors.  
 ACCESSION BC033727  
 VERSION BC033727.1 GI:21707230  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1532)  
 Klausner, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Strausberg, R.L., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.  
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences  
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)  
 2238257  
 12477932  
 2 (bases 1 to 1532)  
 Strausberg, R.  
 Direct Submission  
 Submitted (02-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA  
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
 Contact: MGC help desk  
 Email: [cgabs-r@mail.nih.gov](mailto:cgabs-r@mail.nih.gov)  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),  
 Gaithersburg, Maryland;  
 Web site: <http://www.nisc.nih.gov/>  
 Contact: [nisc\\_mgc@nih.gov](mailto:nisc_mgc@nih.gov)  
 Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAC Plate: 68 Row: i Column: 2  
 This clone was selected for full length sequencing because it passed the following selection criteria: GenomeScan gene prediction  
 This clone has the following problem: frame shifted.  
 Location/Qualifiers  
 1..1532  
 /organism="Homo sapiens"

/mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5166762"  
 /tissue\_type="Brain, adult medulla"  
 /clone\_lib="NIH\_MGC\_l19"  
 /lab\_host="DH10B"  
 /note="Vector: pCMV-SPORT6"

## ORIGIN

Query Match 24.7%; Score 711.2; DB 11; Length 1532;  
 Best Local Similarity 83.3%; Pred. No. 3.3e-109;  
 Matches 910; Conservative 0; Mismatches 8; Indels 174; Gaps 2;  
 QY 378 GTGAGCGCGACCATCGAGTCCGATTATATGTTCTCAAGCGACAGGTGAGAGGTGTTTC 437  
 Db 612 GGGTGGCCACCATGAGAGTCCGATTATATGTTCTCAAGCGACAGGTGAGAGGTGTTTC 671  
 QY 438 GGGCTGAGGAAATACTGTGTCAGTGGTGGCATGGAGTCTCTGGGACACCAAGAT 497  
 Db 672 GGGCTGGAGGAATACTGTGTCAGTGGTGGCATGGAGTCTCTGGGACACCAAGAT 731  
 QY 498 CAGAGGCTACATCCGATACCCAGATTGGCGAAGAACTTCGACGAGGAGCGGTGGCC 557  
 Db 732 CAGAAGGCTACATCCGATACCCATTTTCCGCAAGAACTTCGACGAGGAGCGGTGGCC 791  
 QY 558 AAGGAGGTGTCCCTCGAGCAGGCGCATCGTGTGTCCTGCGCTCCACCGAGGAGTCCCT 617  
 Db 792 AAGGAGGTGTCCCTCGAGCAGGCGCATCGTGTGTCCTGCGCTCCACCGAGGAGTCCCT 851  
 QY 618 CCAGCGAGGTGAGTGGTCCGGAACGAGAGCACTGTGTGGACCCGTCCTTGACCCCAAT 677  
 Db 852 CCAGCGAGGTGAGTGGTCCGGAACGAGAGCACTGTGTGGACCCGTCCTTGACCCCAAT 911  
 QY 678 GTATACATCAGCGGAGCAGACAGCTGTGTGGTGGCAGACGCGCGCTTGTGACACGGCC 737  
 Db 912 GTATACATCAGCGGAGCAGACAGCTGTGTGGTGGCAGACGCGCGCTTGTGACACGGCC 971  
 QY 738 AACTACACCTGCTGGCCAAAGAACATCGTGGCAGCACTGTGGCAGCGCCTCGCTGTGTC 797  
 Db 972 AACTACACCTGCTGGCCAAAGAACATCGTGGCAGCACTGTGGCAGCGCCTCGCTGTGTC 1031  
 QY 798 ATCGTCTACGTGAACGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 857  
 Db 1032 ATCGTCTAC----- 1040  
 QY 858 TGTGGCGCGGTGGCGAGAAACGAGCGGAGTGCACCAACCGCGCGCTCTCAACGGG 917  
 Db 1041 ----- 1040  
 QY 918 GGGCTTTCTGTGAGGGGAGAAATGTCATGACCGCACCGTCTCTCTGTGTTGTCTCT 977  
 Db 1041 ----- 1040  
 QY 978 GTGAGCGGAGCTGGAGCGCGGTGGAGCAAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1037  
 Db 1041 GTGAGCGGAGCTGGAGCGCGGTGGAGCAAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 1100  
 QY 1038 TGGCGGAGCGGTGAGTGTCTGACCCAGACACCCGCAACCGAGGAGGAGTGCAGGGC 1097  
 Db 1101 TGGCGGAGCGGTGAGTGTCTGACCCAGACACCCGCAACCGAGGAGGAGTGCAGGGC 1160  
 QY 1098 ACTGACCTGGACACCGCGAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCT 1157  
 Db 1161 ACTGACCTGGACACCGCGAACTGTACAGTGAACCTCTGTGTACACAGTGTCTTGGCCCT 1220  
 QY 1158 GAGAGCTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTCTGCTGCTGCTGCTGCTGCTG 1217  
 Db 1221 GAGAGCTGGCCCTCTATGTGGGCTCATCGCGTGGCGCTCTGCTGCTGCTGCTGCTGCTG 1280  
 QY 1218 CTTGTCTCTATCTCTCGTTTATTTCGGAAGAAAGAGGGGTGGGACTCAGATGTGGCTGAC 1277  
 Db 1281 CTTGTCTCTATCTCTCGTTTATTTCGGAAGAAAGAGGGGTGGGACTCAGATGTGGCTGAC 1340

```

QY 1278 TCGTCCATTCTACCTCAGGCTTCCAGCCGTCAGCATCAAGCCAGCAAGCAGACAAC 1337
Db 1341 TCGTCCATTCTACCTCAGGCTTCCAGCCGTCAGCATCAAGCCAGCAAGCAGACAAC 1400
QY 1338 CCCATCTGTCTACCATCCAGCCGACCTCAG---CACCAACCACTTACCCAGGCGAGT 1394
Db 1401 CCCATCTGTCTACCATCCAGCCGACCTCAGCACCACCACTTACCCAGGCGAGT 1460
QY 1395 CTCTGTCCCGGCGAGGATGGCCAGCCCAAGCTTCAGCTCAACATGGGCACTGTCTC 1454
Db 1461 CTCTGTCCCGGCGAGGATGGCCAGCCCAAGCTTCAGCTCAACATGGGCACTGTCTC 1520
QY 1455 AGCCCTCTGGGT 1466
Db 1521 AGCCCTCTGGGT 1532

RESULT 12
LOCUS BI818609
DEFINITION 60303362F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5174559 5',
mRNA sequence.
ACCESSION BI818609
VERSION BI818609.1 GI:15929902
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11434 row: 1 column: 16
High quality sequence stop: 744.
FEATURES
    Location/Qualifiers
        1..818
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="IMAGE:5174559"
            /lab_host="DH10B"
            /clone_lib="NIH_MGC_115"
            /note="Organ: pooled brain, lung, testis; Vector:
            pCMV-Sport; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
            source anonymous pool of 6 male brains, age range 23-27; 1
            male lung, age 27; and 1 male testis, age 69. Library is
            oligo-dT primed and directionally cloned (EcoRV site is
            destroyed upon cloning). Average insert size 1.8 kb,
            insert size range 1-3 kb. Library is normalized and
            enriched for full-length clones and was constructed by C.
            Gruber (Invitrogen). Research Genetics tracking code
            021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match 24.6%; Score 710; DB 12; Length 818;
Best Local Similarity 97.1%; Pred. No. 4e-109;
Matches 744; Conservative 0; Mismatches 20; Indels 2; Gaps 2;

QY 84 GCCATGGCCGTCGGCGCGGCTGTGGCCAGCGCTCTGGGCATAGCTCCGCGCTGG 143
Db 1 GCCATGGCCGTCGGCGCGGCTGTGGCCAGCGCTCTGGGCATAGCTCCGCGCTGG 60
QY 144 CTCGCGGCTCGGTGTCGCCAGAGGTGCCACCGTGGCCACCACTGCTGTGTCAC 203

```

```

Db 61 CTCGCGGCTCGGTGTCGCCAGAGAGTGCCACCGTGCCCAACCCAGTGCCTGTGCGCAAC 120
QY 204 CCGGACTGCTTCCCTCCACTTCTCTGTTGGAGCCCGAGGATGTGTACATCGTCAAGACAAG 263
Db 121 CCGGACTGCTTCCCTCCACTTCTCTGTTGGAGCCCGAGGATGTGTACATCGTCAAGACAAG 180
QY 264 CCAAGTCTGCTTGTGTGCAAGCCGTGCCCGCCACACAGATCTTCTTCAAGTCAACGGG 323
Db 181 CCAAGTCTGCTTGTGTGCAAGCCGTGCCCGCCACACAGATCTTCTTCAAGTCAACGGG 240
QY 324 GAGTGGTGGCCAGGTGACACACAGTGTGAGCGGACAGACAGACGGAGCACTGTGTAG 383
Db 241 GAGTGGTGGCCAGGTGACACACAGTGTGAGCGGACAGACAGACGGAGCACTGTGTAG 300
QY 384 CCGACCATGGAGTCCGCAATTAATGCTCAAGCGACAGTGTGAGAGGTTTCGGGCTG 443
Db 301 CCCACCATGGAGTCCGCAATTAATGCTCAAGCGACAGTGTGAGAGGTTTCGGGCTG 360
QY 444 GAGGAATCTGTGTCAGTGCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 503
Db 361 GAGGAATCTGTGTCAGTGCCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 420
QY 504 GCTATCATCCGATAGCAGATTTGCGAAGAACTTCGAGCAGAGCCGTCGCCAAGAG 563
Db 421 GCTATCATCCGATAGCAGATTTGCGAAGAACTTCGAGCAGAGCCGTCGCCAAGAG 480
QY 564 GTGTCCTGGAGCAGCGATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 623
Db 481 GTGTCCTGGAGCAGCGATCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 540
QY 624 GAGTGGAGTGGCTCCGAAACGAGACCTGTGTGGACCCGTCCTGTCGACCCCAATGTATAC 683
Db 541 GAGTGGAGTGGCTCCGAAACGAGACCTGTGTGGACCCGTCCTGTCGACCCCAATGTATAC 600
QY 684 ATCAGCGGGAGCAGACCTGTGTGTGACAGGCGCCGTCGTCGTCGTCGTCGTCGTCGTCG 743
Db 601 ATCAGCGGGAGCAGACCTGTGTGTGACAGGCGCCGTCGTCGTCGTCGTCGTCGTCGTCG 660
QY 744 ACCTGCTGGCCAGCAACATCTGTGACGTC-GCCGACGCGCTCGCTGTCGTCGTCGTCG 802
Db 661 ACCTGCTGGCCAGCAACATCTGTGACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 720
QY 803 CTAGTGAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 847
Db 721 CTAGTGAACGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 766

RESULT 13
EX348193/c
LOCUS BX348193
DEFINITION BX348193 Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED Homo sapiens
cDNA clone CS0DB008YE02 5-PRIME, mRNA sequence.
ACCESSION BX348193
VERSION BX348193.1 GI:30367259
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
1 (bases 1 to 843)
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3239.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF004ZD01_AF00293_1kcluster=3239.r.

```

Contact : Peng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0BAF004ZB01\_AF00293\_1.

## FEATURES

Location/Qualifiers  
1. .843  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DB008YB02"  
/tissue type="NEUROBLASTOMA COT 10-NORMALIZED"  
/clone lib="Homo sapiens NEUROBLASTOMA COT 10-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 24.6%; Score 709.8; DB 13; Length 843;  
Best Local Similarity 95.0%; Pred. No. 4.3e-109;  
Matches 776; Conservative 0; Mismatches 34; Indels 7; Gaps 4;  
QY 1694 CACGCTGCACAGCCGGAAGAGCTGAGG--TTGCCCTTAGTGGCTGTGAGACCCCTGTG 1751  
DB 824 CACGCGGCCAAAGCGGAAGAGTGTAGGTTGCCCAACCCGGGTGTGAGACCCCTGTG 765  
QY 1752 AGTCCATCTTAGTGTGAGACCCCTGTGCTGCTCACCAGCCAGTGTATCTGGCT 1811  
DB 764 AGTCCATCTTAGTGTGAA--CCCTGGCTGCTCACCAGCCAGTGTATCTGGGT 706  
QY 1812 ATGGACCACTGTGGGAGCCAGCCCTGACAGCTGGAGCTGGCCCTC-AAAAAGCAGTC 1870  
DB 705 ATGGACCACTGTGGGAGCCAGCCCTGACAGCTGAGGCTGGCTCATAAAAAAGTC 646  
QY 1871 GTGCGAGGACAGCTGGGAGCAGATGTGTGTACCTTGGGCGAGAGCGCCCTCCACCT 1930  
DB 645 GTGCGAGGACAGCTGGG--AGTATGTCTGTGACCTTGGGCGAGNAGCGCCCTCCACCT 589  
QY 1931 CTACTACTCCAGCTGGAGGCGAGTGTCTGCTACCTTCTTCCAGCAGCTGGCGGCTT 1990  
DB 588 CTACTACTCCAGCTGGAGGCGAGTGTCTGCTACCTTCTTCCAGCAGCTGGCGGCTT 529  
QY 1991 TGCCTGTGGGAGAGGCGCTCAGCTGGCTGGCGCAAGCGCTCAAGCTGCTTCTGT 2050  
DB 528 TGCCTGTGGGAGAGGCGCTCAGCTGGCTGGCGCAAGCGCTCAAGCTGCTTCTGT 469  
QY 2051 TGCCTGTGGCTGCACCTCCTCGAGTACAACTCCGGGTCTATGCTGTGATGACAC 2110  
DB 468 TGCCTGTGGCTGCACCTCCTCGAGTACAACTCCGGGTCTATGCTGTGATGACAC 409  
QY 2111 CCAGATGACTCAAGAGGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2170  
DB 408 CCAGATGACTCAAGAGGTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 349  
QY 2171 GGAGCAGCGGTCTCTGCACTTCAAGGACAGTGTACCAAACTGGCGCTATCCATCCACGA 2230  
DB 348 GGAGCAGCGGTCTCTGCACTTCAAGGACAGTGTACCAAACTGGCGCTATCCATCCACGA 289  
QY 2231 TGTGCCAGCTCCTCTGTGGAAGAGTAAAGTCTTGTGAGTGTGAGTGTGAGTGTGAGT 2290  
DB 288 TGTGCCAGCTCCTCTGTGGAAGAGTAAAGTCTTGTGAGTGTGAGTGTGAGTGTGAGT 229  
QY 2291 TCACATCTGGAATGGCAGCGGTACTTGGCACTGACCTTCCACCTGGAGCGTGTGAG 2350  
DB 228 TCACATCTGGAATGGCAGCGGTACTTGGCACTGACCTTCCACCTGGAGCGTGTGAG 169  
QY 2351 CCCAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2410  
DB 168 CCCAGCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 109  
QY 2411 GAGCTTCAGCATCACTTCAACATCACCAGGACAGAGTTTCTGAGTGTGAGTGTGAGTCT 2470  
DB 108 GAGCTTCAGCATCACTTCAACATCACCAGGACAGAGTTTCTGAGTGTGAGTGTGAGTCT 49

QY 2471 GGAGAGTGAAGCGGGGTCCAGCGCTGTGTTGGGCCCC 2507  
DB 48 GGAGAGTGAAGCGGGGTCCAGCGCTGTGTTGGGCCCC 12  
RESULT 14  
BX345406  
LOCUS  
DEFINITION BX345406 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens  
CDNA clone CS0DC023YB10 5-PRIME, mRNA sequence.  
ACCESSION BX345406  
VERSION BX345406.1 GI:30369001  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Li.W.B., Gruber,C., Jessee,J. and Polayes,D.  
Full-length cDNA libraries and normalization  
Unpublished (2001)  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of  
Invitrogen. This sequence belongs to sequence cluster 3529.f For  
more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CS0AC023DA05QPlcluster=3529.f. Contact :  
Peng Liang Email : fliang@lifetech.com URL :  
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0AC023DA05QPl.  
Location/Qualifiers  
1. .939  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DC023YB10"  
/tissue type="NEUROBLASTOMA COT 25-NORMALIZED"  
/clone lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN  
Query Match 24.3%; Score 700.2; DB 13; Length 939;  
Best Local Similarity 94.0%; Pred. No. 1.9e-107;  
Matches 755; Conservative 22; Mismatches 20; Indels 6; Gaps 5;  
QY 1 AGCTGGGCTCCGGGTGAGGCGCTAAAGCGGCTCCCGCCCGGGGCGGCGGCGGCGG 60  
DB 142 AGCTGGGCTCCGGGTGAGGCGCTAAAGCGGCTCCCGCCCGGGGCGGCGGCGGCGG 201  
QY 61 GCCCGCGGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 120  
DB 202 GCCCGCGGCTCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTC 261  
QY 121 TGGGCTATGCTTCGCGGCTTGGCTCGGCGGCTCGGGTGCAGCAGAGTGCACCGTGG 180  
DB 262 TGGGCTATGCTTCGCGGCTTGGCTCGGCGGCTCGGGTGCAGCAGAGTGCACCGTGG 321  
QY 181 CCAACCCAGTGTGTTGCAACCGGACCTGTTTCCCACTTCTTGGTGGAGCCCGGAG 240  
DB 322 CCAACCCAGTGTGTTGCAACCGGACCTGTTTCCCACTTCTTGGTGGAGCCCGGAG 381  
QY 241 ATGTGTACATCGTCAAGAACCAAGCCAGTGTGTGTGTGCAAGGCGGTGCCCGCACGC 300  
DB 382 ATGTGTACATCGTCAAGAACCAAGCCAGTGTGTGTGTGCAAGGCGGTGCCCGCACGC 439  
QY 301 AGATCTTCTTCAAGTCAACCGGAGTGGGTGGCGCCAGGTGGACACGATGATCGAGCGCA 360



mis Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:10:30 ; Search time 137 Seconds  
(without alignments)  
1854.089 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLNPALLGIVLAAML.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001s.\*
- 5: Geneseq2002s.\*
- 6: Geneseq2003as.\*
- 7: Geneseq2003bs.\*
- 8: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4787	100.0	899	AAU79939	Aau79939 Human UNC
2	4688.5	98.2	898	AAU85403	Aau85403 Human pro
3	4688.5	97.9	898	AAU97899	Aau97899 Human net
4	4545.5	95.0	898	AAW78898	Aaw78898 Rat UNC-5
5	4545.5	95.0	898	AAU10543	Aau10543 Rat netri
6	4545.5	95.0	898	AAU97900	Aau97900 Rat netri
7	4434	92.6	943	AAW79128	Aaw79128 Human pro
8	4379.5	91.5	842	AAW74818	Aaw74818 Human REP
9	2789	58.3	556	AAW78899	Aaw78899 Human UNC
10	2673.5	55.8	931	AAW50691	Aaw50691 Human UNC
11	2673.5	55.8	931	ADW63098	Adw63098 Human pro
12	2673.5	55.8	982	ABG11551	Abg11551 Novel hum
13	2497	52.2	945	ADW63096	Adw63096 Rat Prote
14	2490.5	52.0	946	ABG61795	Abg61795 Novel UNC
15	2490	52.0	943	AAW78900	Aaw78900 Rat UNC-5
16	2484	51.9	933	AAO18734	Aao18734 Human NOV
17	2484	51.9	933	AAO18735	Aao18735 Human NOV
18	2479	51.8	945	AAU12244	Aau12244 Human PRO
19	2479	51.8	945	ABO17688	Abol17688 Novel hum
20	2479	51.8	945	ABU80942	Abu80942 Human PRO
21	2479	51.8	945	ABU66642	Abu66642 Human PRO
22	2479	51.8	945	ABU59723	Abu59723 Novel sec
23	2479	51.8	945	ABO24913	Abol24913 Human sec
24	2479	51.8	945	ABU66918	Abu66918 Human sec
25	2479	51.8	945	ADA45665	Ada45665 Novel hum

## ALIGNMENTS

### RESULT 1

AAU79939  
ID AAU79939 standard; protein; 899 AA.

XX	AC	AAU79939;			
XX	DT	15-JUL-2002 (first entry)			
XX	DE	Human UNC5-like protein NOV1.			
XX	KW	Human; NOVX polypeptide; cardiomyopathy; atherosclerosis; cancer;			
XX	KW	cell signal processing; metabolic pathway modulation; cancerous tissue;			
XX	KW	antibody; diabetes; transgenic animal; UNC5-like protein; NOV1;			
XX	XX	chromosome 13.			
XX	OS	Homo sapiens.			
XX	PN	WO200229038-A2.			
XX	PD	11-APR-2002.			
XX	PF	04-OCT-2001; 2001WO-US031377.			
XX	PR	04-OCT-2000; 2000US-0237862P.			
XX	PA	(CURA-) CURAGEN CORP.			
XX	PI	Heirmann JL, Rastelli L, Shimkets RA;			
XX	DR	WPI: 2002-340104/37.			
XX	DR	N-PSDB; ABK49422.			
XX	PT	Novel isolated NOVX polypeptide, and encoded polynucleotide, useful for			
XX	PS	treating cardiomyopathy, atherosclerosis, and cancer.			
XX	CC	Claim 1; Page 9; 180pp; English.			
XX	CC	The present invention relates to a new NOVX polypeptide having a 900			
XX	CC	(NOV1), 4349 (NOV2), 940 (NOV3), 798 (NOV4), 865 (NOV5), or 311 (NOV6)			
XX	CC	residue amino acid sequence, as given in the specification. The novel			
XX	CC	polypeptide, and its encoding polynucleotide, are used to treat			
XX	CC	cardiomyopathy, atherosclerosis, cancer or a disease related to cell			
XX	CC	signal processing and metabolic pathway modulation, in a human. Detecting			
XX	CC	the polypeptide or polynucleotide is useful for identifying cancerous			
XX	CC	tissue. The antibody can be used to treat diabetes or cancer. The host			
XX	CC	cells can be used to produce non-human transgenic animals useful in drug			
XX	CC	screening. The present amino acid sequence is that of the human UNC5-like			
XX	CC	protein NOV1 of the invention. This sequence is encoded by the human UNC5			

26	2479	51.8	945	6	ADA76096	Ada76096 Human PRO
27	2479	51.8	945	6	ADA18746	Ada18746 Human PRO
28	2479	51.8	945	6	ADA61369	Ada61369 Homo sapi
29	2479	51.8	945	6	ADB19154	Adb19154 Novel hum
30	2479	51.8	945	6	ADB27695	Adb27695 Human PRO
31	2479	51.8	945	6	ADA86174	Ada86174 Novel hum
32	2479	51.8	945	6	ADB15738	Adb15738 Human PRO
33	2479	51.8	945	6	ADA47524	Ada47524 Human PRO
34	2479	51.8	945	6	ADA67319	Ada67319 Human PRO
35	2479	51.8	945	6	ADB30326	Adb30326 Human PRO
36	2479	51.8	945	6	ADA85622	Ada85622 Novel hum
37	2479	51.8	945	6	ADA96834	Ada96834 Human PRO
38	2479	51.8	945	6	ADA79138	Ada79138 Human PRO
39	2479	51.8	945	6	ADB16479	Adb16479 Human PRO
40	2479	51.8	945	6	ADA91571	Ada91571 Novel hum
41	2479	51.8	945	6	ADB14634	Adb14634 Human PRO
42	2479	51.8	945	6	ADB18595	Adb18595 Novel hum
43	2479	51.8	945	6	ADA93810	Ada93810 Human PRO
44	2479	51.8	945	6	ADB19706	Adb19706 Novel hum
45	2479	51.8	945	6		



CC	-like NOV1 gene located on chromosome 13	
XX		
SQ	Sequence 899 AA;	
	Query Match 100.0%; Score 4787; DB 5; Length 899;	
	Best Local Similarity 100.0%; Pred. No. 0;	
	Matches 899; Conservative 0; Mismatches	
QY	1 MAVRGLWPALLGIVLAALWLSGSAQASATVANPVGANPDLLPHFLVEPDVYVKNKP	60
DB	1 MAVRGLWPALLGIVLAALWLSGSAQASATVANPVGANPDLLPHFLVEPDVYVKNKP	60
QY	61 VLLVCKAVPATQIFPKNGEWWVRQVDHVIERSSTDSSGEPTMEVRINVSROQVEKVGLE	120
DB	61 VLLVCKAVPATQIFPKNGEWWVRQVDHVIERSSTDSSGEPTMEVRINVSROQVEKVGLE	120
QY	121 EYWCQVAMSSGTTKSKAYIRIARLRNFKQEPPLAKEVSLQGIIVLPCRPPEGIPPAE	180
DB	121 EYWCQVAMSSGTTKSKAYIRIARLRNFKQEPPLAKEVSLQGIIVLPCRPPEGIPPAE	180
QY	181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY	240
DB	181 VEWLRNEDLVPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVIVY	240
QY	241 VNGGWSWTWMSVCSASGRGWQKRSRSCNTNAPLNGGAFCEGQNVHDTVSSLVSDG	300
DB	241 VNGGWSWTWMSVCSASGRGWQKRSRSCNTNAPLNGGAFCEGQNVHDTVSSLVSDG	300
QY	301 SWSPKWSACGIDCTHWSRECSDPAPRNGEECGTDLDTNRNCTSDLCVHSASGPDV	360
DB	301 SWSPKWSACGIDCTHWSRECSDPAPRNGEECGTDLDTNRNCTSDLCVHSASGPDV	360
QY	361 ALYVGLIAVAVCLVLLVLLVLYCRKXKGLSDVADSSILTSGFQPVSIKPSKADNPHL	420
DB	361 ALYVGLIAVAVCLVLLVLLVLYCRKXKGLSDVADSSILTSGFQPVSIKPSKADNPHL	420
QY	421 LTIQPOLSTTTTYQGSLCPQDQSPKFOITNGHLLSLPLGGGRHTLHSSPTSEAEFVS	480
DB	421 LTIQPOLSTTTTYQGSLCPQDQSPKFOITNGHLLSLPLGGGRHTLHSSPTSEAEFVS	480
QY	481 RLSQNVFSLPRTGNTMYGTENFLGRLMIPNTGILLSIPDAPLPRGIYIYITLHK	540
DB	481 RLSQNVFSLPRTGNTMYGTENFLGRLMIPNTGILLSIPDAPLPRGIYIYITLHK	540
QY	541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCEPSPDSNLSRLKKQSCGWS	600
DB	541 PEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCEPSPDSNLSRLKKQSCGWS	600
QY	601 EQDVLHLGEEAPSHLYYQLEASACVYFTQLGRFALVGEALSVAARKLKLLEAPVAC	660
DB	601 EQDVLHLGEEAPSHLYYQLEASACVYFTQLGRFALVGEALSVAARKLKLLEAPVAC	660
QY	661 TSLEYNIRVYCLDHTDALKEVVOLEKQLGGQLIQPRVLHFKDSYHNLRLSLIHDPSSL	720
DB	661 TSLEYNIRVYCLDHTDALKEVVOLEKQLGGQLIQPRVLHFKDSYHNLRLSLIHDPSSL	720
QY	721 WSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERSVSPSTSLACKLWQVVEGQGFPSIN	780
DB	721 WSKLLVSYQEIIPFYHIWNGTQRYLHCTFTLERSVSPSTSLACKLWQVVEGQGFPSIN	780
QY	781 FNITKTRFAELLALAESEAGVPALVGPSAPKIPFLIRQKILSSLDPPCRRGADWRTLAQK	840
DB	781 FNITKTRFAELLALAESEAGVPALVGPSAPKIPFLIRQKILSSLDPPCRRGADWRTLAQK	840
QY	841 LHLDSHLSFASPKSPPTAMILNWEARHFPNGNLSQAAAAVAGLQGPDAGLFTVSEAE	899
DB	841 LHLDSHLSFASPKSPPTAMILNWEARHFPNGNLSQAAAAVAGLQGPDAGLFTVSEAE	899
RESULT 2	AAU85403 standard; protein; 898 AA.	
ID	AAU85403	
XX		

AC AAU85403;  
XX 21-MAY-2002 (first entry)  
XX Human protein NOV1.  
XX  
XX Human; NOVX; cardiomyopathy; atherosclerosis; diabetes;  
XX cell signal processing disorder; metabolic disorder; obesity; infection;  
XX anorexia; cancer-associated cachexia; cancer; neurodegenerative disorder;  
XX Alzheimer's disease; Parkinson's disease; immune disorder;  
XX haematopoietic disorders; dyslipidaemia; pain; asthma; hypertension;  
XX osteoporosis; Crohn's disease; multiple sclerosis; angina pectoris;  
XX myocardial infarction; ulcer; allergy; benign prostatic hypertrophy;  
XX psychosis; neurological disorder; anxiety; schizophrenia;  
XX manic depression; dementia; dyskinesia; Huntington's disease;  
XX Gilles de la Tourette's syndrome; gene therapy.  
XX  
XX Homo sapiens.  
XX WO200210216-A2.  
XX  
XX 07-FEB-2002.  
XX  
XX 30-JUL-2001; 2001WO-US024225.  
XX  
XX 28-JUL-2000; 2000US-0221409P.  
XX 04-AUG-2000; 2000US-0222840P.  
XX 04-AUG-2000; 2000US-0223752P.  
XX 04-AUG-2000; 2000US-0223762P.  
XX 04-AUG-2000; 2000US-0223769P.  
XX 04-AUG-2000; 2000US-0223770P.  
XX 14-AUG-2000; 2000US-0225146P.  
XX 15-AUG-2000; 2000US-0225392P.  
XX 16-AUG-2000; 2000US-0225697P.  
XX 01-FEB-2001; 2001US-0263662P.  
XX 05-APR-2001; 2001US-0281645P.  
XX (CURA-) CURAGEN CORP.  
XX  
XX Padigar M, Mezes P, Mishra V, Burgess C, Casman S, Grose WM;  
XX Alsobrook JP, Lepley DM, Gerlach VL, Macdougall JR, Smithson G;  
XX  
XX WPI: 2002-180074/23.  
XX N-PSDB; ABK37922.  
XX  
XX New isolated cytoplasmic, nuclear, membrane bound, or secreted  
XX polypeptide, useful for treating cardiomyopathy, atherosclerosis,  
XX infections, cancer, neurodegenerative, metabolic, hematopoietic and  
XX immune disorders.  
XX  
XX Claim 1; Page 11; 213pp; English.  
XX  
XX The invention relates to an isolated cytoplasmic, nuclear, membrane  
XX bound, or secreted polypeptide (NOVX, x= 1-14) their variants or mature  
XX form. Also included are the nucleic acids encoding the NOVX proteins, a  
XX vector comprising the nucleic acid, a cell comprising the vector, an anti  
XX -NOVX antibody and modulators of NOVX. NOVX, the nucleic acid and the  
XX antibody are useful for treating or preventing a NOVX-associated  
XX disorder, where the disorder is selected from cardiomyopathy,  
XX atherosclerosis, diabetes, a disorder related to cell signal processing  
XX and metabolic pathway modulation, metabolic disorders, obesity,  
XX infectious disease, anorexia, cancer-associated cachexia, cancer,  
XX neurodegenerative disorders, Alzheimer's disease, Parkinson's disease,  
XX immune disorders, haematopoietic disorders, and the various  
XX dyslipidaemias, metabolic disturbances associated with obesity, the  
XX metabolic syndrome X and wasting disorders associated with chronic  
XX diseases, bacterial, fungal, protozoal and viral infections, pain,  
XX bulimia, asthma, hypertension, urinary retention, osteoporosis, Crohn's  
XX disease, multiple sclerosis, Albricht Hereditary Osteodystrophy, angina  
XX pectoris, myocardial infarction, ulcer, allergy, benign prostatic  
XX hypertrophy, and psychotic and neurological disorders, including anxiety,  
XX schizophrenia, manic depression, delirium, dementia, and dyskinesias,  
XX



Best Local Similarity 98.4%; Pred. No. 0;		Matches 886; Conservative 3; Mismatches 8; Indels 3; Gaps 3;	
Qy	1	MAVRPGLWPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP	60
Db	1	MAVRPGLWPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP	60
Qy	61	VLLVCVAVPATQIFFKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVFGL	120
Db	61	VLLVCVAVPATQIFFKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVFGL	120
Qy	121	EYWCQCVAWSSSGTTKSKAYIRIARLKNFEQPLAKEVSLQGIPLCPRPPEGIPPAE	180
Db	121	EYWCQCVAWSSSGTTKSKAYIRIARLKNFEQPLAKEVSLQGIPLCPRPPEGIPPAE	180
Qy	181	VEWLNRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Db	181	VEWLNRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240
Qy	241	VNGGWSTWTEWMSVCSASGRGHQKRSRSCNTNAPLNGGAFCEGQNV-KTACATLCPVDG	300
Db	241	VNGGWSTWTEWMSVCSASGRGHQKRSRSCNTNAPLNGGAFCEGQNV-KTACATLCPVDG	299
Qy	301	SWSPKSKWSACGLDCTHWRSCSDPAPRNGEBCQGTDLTRNCTSDLCVHSASGPEV	360
Db	300	SWSPKSKWSACGLDCTHWRSCSDPAPRNGEBCQGTDLTRNCTSDLCVHTASGPEV	359
Qy	361	ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	420
Db	360	ALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSILTSGFQPVSIKPSKADNPHL	419
Qy	421	LTIQPDLS-TTTTYQSLCPRQDPSKQTLNGLHLLSPGGRHTLHSSSTSEAEFV	479
Db	420	LTIQPDLS-TTTTYQSLCPRQDPSKQTLNGLHLLSPGGRHTLHSSSTSEAEFV	479
Qy	480	SRISTQNYFRSLPRGTSNMITYGFNFGRLMIPNTGISILLIPDDAI PRGKIYIYLTILH	539
Db	480	SRISTQNYFRSLPRGTSNMITYGFNFGRLMIPNTGISILLIPDDAI PRGKIYIYLTILH	539
Qy	540	KPDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDRHCGEPPSPDSWSRLKKQCEGS	599
Db	540	KPDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDRHCGEPPSPDSWSRLKKQCEGS	599
Qy	600	WEQDLVHLGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAAAKELKLLFAPVA	659
Db	600	WE-DVLHLGEAPSHLYYCOLEASACVVFTEQLGRFALVGEALSVAAAKELKLLFAPVA	658
Qy	660	CTSLVYNIYVCLDHTDALKEVVQLEKQGGQIOEPVILHFKDSYHNLRLSIHDVPSS	719
Db	659	CTSLVYNIYVCLDHTDALKEVVQLEKQGGQIOEPVILHFKDSYHNLRLSIHDVPSS	718
Qy	720	LWKSLLVSYQEIIPFYHNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVGGQSF	779
Db	719	LWKSLLVSYQEIIPFYHNGTQRYLHCTFTLERVSPSTSDLACKLWVWQVGGQSF	778
Qy	780	NFNITKDTREAEILALESEAGVPAVCPSAFKIPFLIROKIISSLDPPCRGADWRTLAQ	839
Db	779	NFNITKDTREAEILALESEAGVPAVCPSAFKIPFLIROKIISSLDPPCRGADWRTLAQ	838
Qy	840	KLHLDLSLSPFASKPSTMTLNLWEARHPNGLSQLAAVAGLQPDAGLFTVSEAC	899
Db	839	KLHLDLSLSPFASKPSTMTLNLWEARHPNGLSQLAAVAGLQPDAGLFTVSEAC	898
RESULT 4			
ID	AAW78898 standard; protein; 898 AA.		
XX	AAW78898;		
AC	(revised)		
DT	25-MAR-2003 (first entry)		
DT	21-DEC-1998 (first entry)		
XX			

DE	Rat UNC-5 homologue UNC5H-1.		
XX	UNC-5; UNC5H-1; rat; netrin receptor; cell migration; axon guidance;		
KW	diagnosis; therapy.		
OS	Rattus sp.		
XX	Key Location/Qualifiers		
FT	Peptide 580..594		
XX	/note= "peptide used to raise rabbit polyclonal antiser"		
FN	WO9837085-A1.		
XX	27-AUG-1998.		
XX	19-FEB-1998; 98WO-US003143.		
XX	19-FEB-1997; 97US-00808982.		
XX	(REGC ) UNIV CALIFORNIA.		
PA	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;		
PI	WPI; 1998-495364/42.		
DR	N-PSDB; AAV52940.		
XX	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and		
PT	the biopharmaceutical industry.		
XX	Claim 1; Page 19-22; 32pp; English.		
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis		
CC	elegans UNC-5 protein. Their amino acid sequences were deduced from		
CC	isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an		
CC	E18 brain cDNA library. The predicted proteins show similarity with UNC-		
CC	5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin		
CC	type-1 repeats, a predicted membrane spanning region, and a large		
CC	intracellular domain. They are predicted to be involved in cell migration		
CC	and axon guidance, and are characterised as receptor proteins for		
CC	netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins		
CC	are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly		
CC	from transfected host cells. The invention also provides unc-5		
CC	hybridisation probes and primers, vertebrate UNC-5-specific binding		
CC	agents such as specific antibodies, and methods of making and using the		
CC	subject compositions in diagnosis (e.g. genetic hybridisation screens for		
CC	vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate		
CC	vertebrate unc-5 gene expression) and in the biopharmaceutical industry		
CC	(e.g. as immunogens, reagents for modulating cell guidance, reagents for		
CC	screening chemical libraries for lead pharmacological agents, etc.).		
CC	(Updated on 25-MAR-2003 to correct PI field.)		
XX	Sequence 898 AA;		
Qy	Query Match 95.0%; Score 4545.5; DB 2; Length 898;		
Db	Best Local Similarity 94.7%; Pred. No. 0;		
XX	Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;		
Qy	1 MAVRPGLPALLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP	60	
Db	1 MAVRPGLPVLLGIVLAALRGSGAQOQATVANPVGANPDLLPHFLVEPEDVYVKNKP	60	
Qy	61 VLLVCVAVPATQIFFKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVFGL	120	
Db	61 VLLVCVAVPATQIFFKNGEWRQVDHVIERTDSSGSEPTMEVNRINVSQQVEKVFGL	120	
Qy	121 EYWCQCVAWSSSGTTKSKAYIRIARLKNFEQPLAKEVSLQGIPLCPRPPEGIPPAE	180	
Db	121 EYWCQCVAWSSSGTTKSKAYIRIARLKNFEQPLAKEVSLQGIPLCPRPPEGIPPAE	180	
Qy	181 VEWLNRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240	
Db	181 VEWLNRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY	240	

QY	241	VNGGWSTWTSVCSACSGRWQKRSRSCCTNPAPINGAGFCEGQNVDHRTVSLLIVSDG	300
Dd	241	VNGGWSTWTSVCSACSGRWQKRSRSCCTNPAPINGAGFCEGQNVDHRTVSLLIVSDG	299
QY	301	SNSPMSKWSACGLDCTHWRSRECSDDPAPRNGBECOGTDLDRNCTSDLCVHSASGPEDV	360
Dd	300	SWSWSKWSACGLDCTHWRSRECSDDPAPRNGBECRGAOLDTRNCTSDCLHLTASCPEDV	359
QY	361	ALYVGLIYAVAVCLLVLLLVLIVYCRKKGLSDVDASSILTSGFOPVSIKPSKADNPHL	420
Dd	360	ALYIGLVAVAVCLFLLLALLGLIYCRKKGLSDVDASSILTSGFOPVSIKPSKADNPHL	419
QY	421	LTIQPDL- TTTTYQGSCLPRDGSPKFQLTNHLLSPLGGGRHTLHSSPTSEAEEVF	479
Dd	420	LTIQPDLSTTTTTYQGSCLSRDGDSPKFQLSNHLLSPLGSGRHTLHSSPTSEAEEVF	479
QY	480	SLSTQNYFRSLPRGTSNMTYTGNPLGRLMIPNTGISLLIPDAIPRGKIYEYLTLH	539
Dd	480	SLSTQNYFRSLPRGTSNMAYGTGNPLGRLMIPNTGISLLIPDAIPRGKIYEYLTLH	539
QY	540	KPEDVRLPLAGCOTILLSPVSCGPGVLLTRPVILLAMDHCCGPPSPDSWSLRKLKQSCBGS	599
Dd	540	KPEDVRLPLAGCQTLLSPVVSCGPGVLLTRPVILLAMDHCCGPPSPDSWSLRKLKQSCBGS	599
QY	600	WEQDYLHGEEAPSRLHYCYQLEASACYFTQELGRFALVGEALSVAALKRLLLPAPVA	659
Dd	600	WE-DVYLHGLESPSHLYCYQLEAGACYFTQELGRFALVGEALSVAATKRLULLLPAPVA	658
QY	660	CTSLEYNIRVCLHDTHDALKEVVOLEKOLQGQLIOEPRVLHFKDSYHNLRLSIHDPVSS	719
Dd	659	CTSUEYNIRVCLHDTHDALKEVVOLEKOLQGQLIOEPRVLHFKDSYHNLRLSIHDPVSS	718
QY	720	LWKSLLYSYOBIPIYHIWNGTORVYLHCTFTLERVSPSTSDLAACKLWWQVEGDCQSFSI	779
Dd	719	LWKSLLYSYOBIPIYHIWNGTOQQVYLHCTFTLERINASTDLACKVWWQVEGDCQSFSNI	778
QY	780	NFNITKDTRFALLALESEAGVPALVGPSAFKIPFLIRQKIIISLDPPPCCRGAOWRTLQAQ	839
Dd	779	NFNITKDTRFALLALESEGGVPALVGPSAFKIPFLIRQKIIISLDPPCSRGAOWRTLQAQ	838
QY	840	KLHLDSHLSFPASPSPPTAMILNLWEARHPNGMLSQLAAAVAGLGQPDAGLFTVSEAEC	899
Dd	839	KLHLDSHLSFPASPSPPTAMILNLWEARHPNGMLGQLAAAVAGLGQPDAGLFTVSEAEC	898
RESULT 5			
AAU10543			
ID	AAU10543	standard; protein; 898 AA.	
XX	AAU10543;		
AC	XX		
DT	XX	(first entry)	
DE	XX	Rat netrin receptor UNC5H1 (YSG7) polypeptide.	
KW	XX	YSG7; schizophrenia; chronic animal model; LCGU; netrin receptor;	
KW	XX	local cerebral glucose utilisation; phosphodiesterase I-alpha; UNC5H1;	
KW	XX	calcium-independent alpha-latrotoxin receptor; CIRL; trkE; synapsin 1A;	
KW	XX	epithelial discoidin domain receptor 1; synapsin 1B; neuroleptic;	
KW	XX	tumour necrosis factor alpha; TNF-alpha; rat.	
OS	XX	Rattus sp.	
PN	XX	WO200175440-A2.	
PD	XX	11-OCT-2001.	
PF	XX	02-APR-2001; 2001WO-GB001486.	
PR	XX	31-MAR-2000; 2000GB-00007880.	
XX	XX	26-MAY-2000; 2000GB-00012768.	
PA		(WELF-) WELFIDE CORP.	

QY 600 WEQDVLHGEAPSHLYYQLEASACVYFTQGRFALVGEALSVAAAKRLKLLFAPVA 659  
DB 600 WE-DVLHGEAPSHLYYQLEAGACVYFTQGRFALVGEALSVAAATKRLKLLFAPVA 658  
QY 660 CTSLEYNIRVYCLDTHDALKEVQLEKOLGQQLIQEPRVLHFKDSYHNRLSLHDVPS 719  
DB 659 CTSLEYNIRVYCLDTHDALKEVQLEKOLGQQLIQEPRVLHFKDSYHNRLSLHDVPS 718  
QY 720 LKWSKLLVSYQEIPIYHFWNGTQRYLHCTFTLERSVSTSDLACKLWYQVGEQGQSF 779  
DB 719 LKWSKLLVSYQEIPIYHFWNGTQRYLHCTFTLERSVSTSDLACKLWYQVGEQGQSF 778  
QY 780 NFNITKTRFAELLALESEAGVPALVGPFAKIPFLIRKQIISLDDPCRCGADWRTLAQ 839  
DB 779 NFNITKTRFAELLALESEAGVPALVGPFAKIPFLIRKQIISLDDPCRCGADWRTLAQ 838  
QY 840 KLHLDLHLSFPAKSPKPTAMILNWEARHPFNGLSOLAAVAGLGQPDAGLFTVSEAC 899  
DB 839 KLHLDLHLSFPAKSPKPTAMILNWEARHPFNGLSOLAAVAGLGQPDAGLFTVSEAC 898

RESULT 6  
AAU97900  
ID AAU97900 standard; protein; 898 AA.

XX AAU97900;  
DT 27-AUG-2002 (first entry)

XX Rat netrin binding membrane receptor UNC5H-1 protein.

XX Netrin binding membrane receptor; receptor; UNC5H-1; Rat; neurotropic;  
KW neuroprotective; cytosolic; antiparkinsonian; cerebroprotective; cancer;  
KW central nervous system; CNS; stroke; Parkinson's disease;  
KW multiple sclerosis; Alzheimer's disease.

XX Rattus sp.  
XX Key Location/Qualifiers  
FH Domain /note= "Immunoglobulin domain "  
FT Domain /note= "Thrombospondine type 1 domain "  
FT Domain /note= "Thrombospondine type 1 domain "  
FT Region /note= "Transmembrane region"  
FT Domain /note= "ZU5 domain"  
FT Domain /note= "Death domain"

XX W0200233080-A2.  
XX 25-APR-2002.  
XX 15-OCT-2001; 2001WO-EF011891.  
XX 16-OCT-2000; 2000US-0240061P.

XX (FARB ) BAYER AG.  
XX Koehler RH;  
XX WPI; 2002-463314/49.

XX Novel human netrin binding membrane receptor polypeptide and  
PT polynucleotides for identifying modulating agents useful in treating  
PT diseases e.g. Parkinson's disease, multiple sclerosis, stroke,  
PT Alzheimer's disease.  
XX Disclosure; Fig 3; 94pp; English.

XX This invention relates to the DNA and protein sequences of a novel  
CC purified human netrin binding membrane receptor, UNC5H-1. The DNA  
CC sequence of the invention is useful as a probe for detecting a nucleic  
CC acid encoding the UNC5H-1 protein in a biological sample. The sequences  
CC of the invention are useful to screen for agents which decrease the  
CC activity of the UNC5H-1 protein. The sequences are also useful for  
CC screening agents which regulate (modulate) the activity of the protein of  
CC the invention. A pharmaceutical composition containing the protein of the  
CC invention or a reagent that modulates the activity of the UNC5H-1 protein  
CC may be useful for treating a UNC5H-1 dysfunction related disease such as  
CC cancer or a central nervous system (CNS) disorders (e.g. Parkinson's  
CC disease, multiple sclerosis, stroke and Alzheimer's disease). Fusion  
CC proteins comprising the UNC5H-1 protein are useful for generating  
CC antibodies and for in various assay systems, and the protein can be used  
CC as a bait protein in a two-hybrid assay or three-hybrid assay. The method  
CC of the invention is useful for detecting a coding sequence for the UNC5H-  
CC 1 protein. The present sequence represents the Rat netrin binding  
CC membrane receptor UNC5H-1 protein of the invention  
XX Sequence 898 AA;

Query Match 95.0%; Score 4545.5; DB 5; Length 898;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPLMPALIGIVLAAWLRGSGAQASATVANPVGANPDLLPHFLVEPEDVIYVKNKP 60  
DB 1 MAVRPLMPVLLIGIVLAAWLRGSGAQASATVANPVGANPDLLPHFLVEPEDVIYVKNKP 60  
QY 61 VLLVCKKAVPATQIFFKCGEYWRQVDHVIERTSDSSGSEPTMEVRIINVSROQKVEKFGLE 120  
DB 61 VLLVCKKAVPATQIFFKCGEYWRQVDHVIERTSDSSGSEPTMEVRIINVSROQKVEKFGLE 120  
QY 121 EYWCQVAVSSGTTKSKAYIRIARLNKFEQELAKEVSLQEGVILPCRPPEGIPPAE 180  
DB 121 EYWCQVAVSSGTTKSKAYIRIARLNKFEQELAKEVSLQEGVILPCRPPEGIPPAE 180  
QY 181 VEWLNRNEDLVDPDNPVITREHSLVVRQARLADTANTYCAKNIVARRRSASAAVIVY 240  
DB 181 VEWLNRNEDLVDPDNPVITREHSLVVRQARLADTANTYCAKNIVARRRSASAAVIVY 240  
QY 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSLVSDG 300  
DB 241 VNGGWSWTWESVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSLVSDG 299  
QY 301 SWSFWSKWSACGLDCTHWSRECSDPAPRNGGECOCGTDLDRNCTSDLCVHSASGPEDV 360  
DB 300 SWSFWSKWSACGLDCTHWSRECSDPAPRNGGECOCGTDLDRNCTSDLCVHSASGPEDV 359  
QY 361 ALYVGLIYAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQFVSIKPSKADNPHL 420  
DB 360 ALYIGLVAVAVCLVLLVLLVLYCRKKEGLSDVADSSILTSFGQFVSIKPSKADNPHL 419  
QY 421 LTIQPDLS-TTTTYQGSCLPRQDGPSPKFOLTNHLLSLPLGGGRHTLHSSPTSEAEFV 479  
DB 420 LTIQPDLS-TTTTYQGSCLPRQDGPSPKFOLSNHLLSLPLGGGRHTLHSSPTSEAEFV 479  
QY 480 SRLSTQNYFRSLPRGTSNMTYGTNFIIGRGLMIPNTGISLILIPDAIPRGIYEIYTLH 539  
DB 480 SRLSTQNYFRSLPRGTSNMTYGTNFIIGRGLMIPNTGISLILIPDAIPRGIYEIYTLH 539  
QY 540 KPEDVRLPLAGCOTLLSPIVSCGPPGVLLTRPVLAMDHCGEFSPDSWSURLKKQSCGS 599  
DB 540 KPEDVRLPLAGCOTLLSPIVSCGPPGVLLTRPVLAMDHCGEFSPDSWSURLKKQSCGS 599  
QY 600 WEQDVLHGEAPSHLYYQLEASACVYFTQGRFALVGEALSVAAAKRLKLLFAPVA 659  
DB 600 WE-DVLHGEAPSHLYYQLEAGACVYFTQGRFALVGEALSVAAATKRLKLLFAPVA 658  
QY 660 CTSLEYNIRVYCLDTHDALKEVQLEKOLGQQLIQEPRVLHFKDSYHNRLSLHDVPS 719  
DB 659 CTSLEYNIRVYCLDTHDALKEVQLEKOLGQQLIQEPRVLHFKDSYHNRLSLHDVPS 718



AAU74818  
ID AAU74818 standard; protein; 842 AA.  
AC AAU74818;  
XX  
XX 23-APR-2002 (first entry)  
XX Human REPTR 1 protein.  
XX  
XX REPTR: human; antiinflammatory; cytostatic; immunosuppressive; antiviral;  
KW anti-HIV; antiarthritic; anticonvulsant; nootropic; neuroprotective;  
KW antiallergic; antibody; immunogen; endometriosis;  
KW gastrointestinal disorder; gastritis; oesophageal carcinoma;  
KW Crohn's disease; irritable bowel syndrome; ulcerative colitis;  
KW endocrine disease; hypothyroidism; Kallman's disease;  
KW autoimmune disease; inflammatory disease; infertility; receptor;  
KW acquired immune deficiency syndrome; AIDS; rheumatoid arthritis; allergy;  
KW osteoarthritis; diabetes mellitus; multiple sclerosis;  
KW systemic lupus erythematosus; cell proliferative disorder; cancer;  
KW developmental disorder; Duchenne muscular dystrophy;  
KW Becker muscular dystrophy; neurological disorder; epilepsy;  
KW Alzheimer's disease; Huntington's disease; reproductive disorder.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO2001198354-A2.  
XX 27-DEC-2001.  
XX  
XX 21-JUN-2001; 2001WO-US019942.  
XX  
XX 21-JUN-2000; 2000US-0214027P.  
PR 25-AUG-2000; 2000US-0228045P.  
PR 12-DEC-2000; 2000US-0255104P.  
XX  
XX (INCY) INCYTE GENOMICS INC.  
XX  
XX Griffin JA, Kallick DA, Tribouley CM, Yue H, Nguyen DB, Tang YT;  
PI Lal P, Policky JL, Azimzai Y, Lu DAM, Graul R, Yao MG, Burford N;  
PI Hafalia AJA, Baughn MR, Bhandan O, Patterson C, Yang J, Xu Y;  
PI Gandhi AR, Warren BA, Ding L, Sanjanwala MS, Duggan BM, Lu Y;  
XX  
XX WPI; 2002-090432/12.  
DR N-PSDB; ABK15169.  
XX  
XX Twelve human receptors (referred to as REPTR-1 to REPTR-12), useful in  
PT the diagnosis, treatment and prevention of gastrointestinal (e.g.  
PT gastritis), autoimmune/inflammatory (e.g. osteoarthritis) and cell  
PT proliferative (e.g. cancer) disorders.  
XX  
XX Claim 45; Page 111-113; 157pp; English.  
XX  
XX This invention relates to twelve human receptors cDNA sequences referred  
CC to as REPTR-1 to REPTR-12, and the proteins encoded thereby. The  
CC proteins of the invention may have antiinflammatory, cytostatic,  
CC immunosuppressive, antiviral, anti-HIV, antiarthritic, muscular active  
CC general, anticonvulsant, nootropic, neuroprotective, antiallergic  
CC agonists or antagonists, and the protein sequences may be used to raise  
CC anti-REPTR antibodies. These molecules and the REPTR polynucleotides and  
CC polypeptides of the invention are useful in the diagnosis, treatment and  
CC prevention of gastrointestinal (e.g. gastritis, oesophageal carcinoma,  
CC Crohn's disease, irritable bowel syndrome, ulcerative colitis), endocrine  
CC (e.g. hypothyroidism, Kallman's disease), autoimmune/ inflammatory  
CC (e.g. acquired immune deficiency syndrome (AIDS), rheumatoid arthritis,  
CC allergies, osteoarthritis, diabetes mellitus, multiple sclerosis,  
CC systemic lupus erythematosus), cell proliferative (e.g. cancer),  
CC developmental (e.g. Duchenne and Becker muscular dystrophy), neurological  
CC (e.g. epilepsy, Alzheimer's disease, Huntington's disease) and  
CC reproductive (e.g. infertility, endometriosis) disorders. Numerous other  
CC examples of each disorder are given in the specification. The present  
CC sequence represents the human REPTR1 protein sequence of the invention  
XX

SQ	Sequence 842 AA;	
Query Match	91.5%;	Score 4379.5; DB 5; Length 842;
Best Local Similarity	93.1%;	Pred. No. 0;
Matches 838; Conservative	1;	Mismatches 2; Indels 59; Gaps 3;
QY	1	MAVRPGLWPALLGIVLAALRSGAQQSATVANPVGANPDLLPHFLVEPDDVYIVKNKP 60
DB	1	MAVRPGLWPALLGIVLAALRSGAQQSATVANPVGANPDLLPHFLVEPDDVYIVKNKP 60
QY	61	VLLVCVAVPATOIPFKCNQGEVVRQVDHVIERTDSSGCEPTMEVRINVSQQQVEKVGLE 120
DB	61	VLLVCVAVPATOIPFKCNQGEVVRQVDHVIERTDSSGCEPTMEVRINVSQQQVEKVGLE 120
QY	121	EYWCOCVAVSSGTTKSKAVIRIARLNKFEQELAKEVSLQEGIVLPCRPPEGIPPAE 180
DB	121	EYWCOCVAVSSGTTKSKAVIRIARLNKFEQELAKEVSLQEGIVLPCRPPEGIPPAE 180
QY	181	VEWLRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
DB	181	VEWLRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY	241	VNGGWSWTWESVCSACGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300
DB	241	-----VDG 243
QY	301	SWSPWSKWSACGLDCTHWRSECSDPAPRNGGEECGQTDLTDRNCTSDLCVHSASGPDV 360
DB	244	SWSPWSKWSACGLDCTHWRSECSDPAPRNGGEECGQTDLTDRNCTSDLCVHTASGPDV 303
QY	361	ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSGFQPVSKSKADNPHL 420
DB	304	ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTSGFQPVSKSKADNPHL 363
QY	421	LTIQDPLS-TTTTYQGLCPRODGSPPKQFQNLNGHLLSPGGRHLLHSSPTSEAEFV 479
DB	364	LTIQDPLSTTTTYQGLCPRODGSPPKQFQNLNGHLLSPGGRHLLHSSPTSEAEFV 423
QY	480	SRLSTQNYPRSLPRGTSNMVYTFNPLGRLMIPNTGISLLIPPDAPRGIYEYLTLLH 539
DB	424	SRLSTQNYPRSLPRGTSNMVYTFNPLGRLMIPNTGISLLIPPDAPRGIYEYLTLLH 483
QY	540	KPEDVRLPLAGCQTLSPISVSCGPGVLLTRPVILAMDHCGBPSPDSWSLRLKKQCEGS 599
DB	484	KPEDVRLPLAGCQTLSPISVSCGPGVLLTRPVILAMDHCGBPSPDSWSLRLKKQCEGS 543
QY	600	WEQDVLHLGEBAPSHLYYQLEASACYVTEQIGREALVGEALSVAAAKELKLLLPAPVA 659
DB	544	WE-DVHLGEBAPSHLYYQLEASACYVTEQIGREALVGEALSVAAAKELKLLLPAPVA 602
QY	660	CTSLEYNIRVYCLHDTHDALKKEVQLEKOLGGQLIQEPVRLHFKDSYHNLRLSIHDVPSS 719
DB	603	CTSLEYNIRVYCLHDTHDALKKEVQLEKOLGGQLIQEPVRLHFKDSYHNLRLSIHDVPSS 662
QY	720	LMKSKLLVSYQEIPIFYHINWGTORYLHCTFTLRRVSPSTSDACKLWVQVEGDSQFSI 779
DB	663	LMKSKLLVSYQEIPIFYHINWGTORYLHCTFTLRRVSPSTSDACKLWVQVEGDSQFSI 722
QY	780	NFNITKDTFAELALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
DB	723	NFNITKDTFAELALAESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 782
QY	840	KLHLDHSLSPFASKPSPTAMINLWEARHFPNGNLSQAAAAGLQDPDAGLFTVSEAC 899
DB	783	KLHLDHSLSPFASKPSPTAMINLWEARHFPNGNLSQAAAAGLQDPDAGLFTVSEAC 842
RESULT 9		
AAW78899		
ID	AAW78899	standard; protein; 556 AA.
XX	AC	AAW78899;
XX	XX	





identifying unknown human cDNAs which encode proteins that interact with the human unc-5C protein.

Disclosure; Page 224-227; 246pp; English.

The present invention describes 3 variants of human unc-5C cDNAs (unc-5Cb, unc-5Cc and unc-5C8) which correspond to alternatively spliced unc-5C transcripts, and a human unc-5HS1 cDNA which shares homology with the Rattus norvegicus unc-5HS1 cDNA. Also described are assays based on protein-protein-interactions between the unc-5 protein and a variety of different interacting proteins. The unc-5C variant cDNAs and unc-5HS1 cDNA are useful in methods for identifying compounds which reduce or inhibit the lethal phenotype associated with the expression of the unc-5 death domain in yeast. They are also useful in yeast two hybrid experiments for identifying unknown human cDNAs which encode proteins that interact with the human unc-5C protein. AAC90914 to AAC90971 and AAC90646 to AAC90693 represent sequences used in the exemplification of the present invention

Sequence 931 AA;

Query Match 55.8%; Score 2673.5; DB 4; Length 931;  
Best Local Similarity 55.5%; Pred. No. 1.1e-222;  
Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;

9 FALLGIVLAAMLRGSAQOS---ATVANPVGANFDLLPHFLVPEDEVIIVKPKVLLVC 65  
26 PAL--ALLSASGTGAAQDDFFHELFPETFPDPPEPLPHFLIEPEAVIVKPKVNLVC 83  
66 KAVPATQIFKNGEVRQVDHVIERS TDGSGEPTMEVRINVSROQVEKVFGLREYWCQ 125  
84 KASPATQIFKNGEVRQVDHVIERS TDGSGEPTMEVRINVSROQVEKVFGLREYWCQ 143  
126 CVAWSSGTTKOKAYIRIARLRKNEQEPFLAKVNSLEQGVLPKCRPPGIPPAEVEVLR 185  
144 CVAWSSGTTKOKAYIRIARLRKNEQEPFLAKVNSLEQGVLPKCRPPGIPPAEVEVLR 203  
186 NEDVDPSPDNPVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAAVIVYNGW 245  
204 NEDVDPSPDNPVYITREHSLVVRQARLADTANYTCVAKNIVARRSASAAVIVYNGW 263  
246 STWTEWSVCSACRGWOKRSCTNPALNGAFCEGQVVDRTVSSLLVSDGGSWP 305  
264 STWTEWSVCSACRGWOKRSCTNPALNGAFCEGQVVDRTVSSLLVSDGGSWP 322  
306 SKWSACGLDCTHWRSECDPAPNGEGECQGTDLDRNCTSDLCVHSASGPEDDVALYVG 365  
323 SKWSTCGTECHWRRECTAPAPNGGKDCDGLVLSQKNTDGLCMTAPSDSDVALYVG 382  
366 L-TAVAVCLVLLLLVILVYCRKEGSDVDASSILTSFGFQVSIKPSKADNPHLLTIQ 424  
383 IVIAVIVCLAISSVVALFVYKRNHRDFESDIIDSSALNGFGFQVNIKAARQD---LLAVP 439  
425 PDL-SITTTTQGSICPRQDQPSKFQLTNGHLLSPGCGRHHTLHS-----SPTSEAEFV 479  
440 PDLTSAAMVYRGPVYALHD-VSDKIPMTNSPILDPILNLIKVINTSAGVSPQDDISEFT 498  
480 SRLS---TONYF-----RSLPRGT--SNMTYGTFFNLGGLRMTINTGILLSLIPDDA 525  
499 SKLSPQMTQSLLENEALSLSKQSLARQTDFTSCTAFGSFNSLGGHLIVPNSGVSLIPAGA 558  
526 IPKGYIYILTHKPEDVRLPLAGCQTLSPISVPCGPPGVLLTRPVILAMDHGCRSPD 585  
559 IPQGRVYVYVTHRKETMRPMDSDQTLTTPVVSQCPGALLTRPVVLTMHHCADNTE 618  
586 SWSLRKQSCGSEWQDVLHLEAPSHLYCYOLEASACVYFTEQLGREALVGEALSVA 645  
619 DWKILLKNQAQGGWE-DVVVVGENTTPCYIKLDEACHILTENLSVALVGHSTTKA 677  
646 AAKELKLLFAPVACTSLEYNIRYVCLHDTDALKEVVQLEKQGGQIIGEPVILHPKDS 705  
678 AAKELKLAIFGLCCSSLEYSIRVYCLDDTDQALKEILHLERQTGGQLLEPKALHPKGS 737

706 YHNRLSLHDPVSSILWKSLLVSYQEIPIFYHINWGTQRYLHCTFTLBRVSPSTSLACKL 765  
738 THNRLSLHDIASHLSKSLKLLAKYQEIPIFYHVWSSQRLHCTFTLBRFSLNVELVCKL 797  
766 WVVQVEGQGSFNFNITKDTFRFAELLALSEAGVPAALVGPSPAFKIPFLIROKLISSLD 825  
798 CVRQVEGEGQIFQLNCTVSEBPTGIDPLDLPANTITVTGFSAPFIPFLIROKLISSLD 857  
826 PPCRGADWRTLAQKLHLDLSHLSFASKPSPTAMTLNWEARHPNGNLSOLAAAVAGLG 885  
858 APQTRGDWRMLAKHLNDRYLNIFATKSSPTGVLIDLWEAQNFDPDGNLSMLAAVLEWVG 917  
886 QPDAGLFTVSEAE 898  
918 RHEIVVSLAERGQ 930  
RESULT 11  
ADE63098  
ID ADE63098 standard; protein; 931 AA.  
XX AC ADE63098;  
XX DT 29-JAN-2004 (first entry)  
XX Human Protein AAC67491, SEQ ID NO 9033.  
DE Human; pain; neuronal tissue; gene therapy;  
KW spinal segmental nerve injury; chronic constriction injury; CCI;  
KW spared nerve injury; SNI; Chung.  
OS Homo sapiens.  
XX WO2003016475-A2.  
XX 27-FEB-2003.  
XX 14-AUG-2002; 2002WO-US025765.  
XX 14-AUG-2001; 2001US-0312147P.  
XX 01-NOV-2001; 2001US-0346382P.  
XX 26-NOV-2001; 2001US-0333347P.  
XX (GEO) GEN HOSPITAL CORP.  
XX (FARB) BAYER AG.  
XX Woolf C, D'urso D, Befort K, Costigan M;  
XX WPI; 2003-268312/26.  
XX GENBANK; AAC67491.  
XX New composition comprising two or more isolated polypeptides, useful for  
XX preparing a medicament for treating pain in an animal.  
XX Claim 1; Page; 1017pp; English.  
XX The invention discloses a composition comprising two or more isolated rat  
XX or human polynucleotides or a polynucleotide which represents a fragment,  
XX derivative or allelic variation of the nucleic acid sequence. Also  
XX claimed are a vector comprising the novel polynucleotide, a host cell  
XX comprising the vector, a method for identifying a nucleotide sequence  
XX which is differentially regulated in an animal subjected to pain and a  
XX kit to perform the method, an array, a method for identifying an agent  
XX that increases or decreases the expression of the polynucleotide sequence  
XX that is differentially expressed in neuronal tissue of a first animal  
XX subjected to pain, a method for identifying a compound which regulates  
XX the expression of a polynucleotide sequence which is differentially  
XX expressed in an animal subjected to pain, a method for identifying a  
XX compound that regulates the activity of one or more of the  
XX polynucleotides, a method for producing a pharmaceutical composition, a  
XX method for identifying a compound or small molecule that regulates the  
XX activity in an animal of one or more of the polypeptides given in the  
XX specification, a method for identifying a compound useful in treating







Best Local Similarity 52.1%; Pred. No. 9.7e-207;		Matches 496; Conservative 150; Mismatches 247; Indels 59; Gaps 18;	
QY	1	MAVRPGLWPA	11
DB	1	NGARSGARGALL	60
QY	57	KNKPVLLVCKAV	111
DB	61	KNKPVLLVCKAV	120
QY	112	QVEKVFGLLEW	171
DB	121	QVEELFGLLEW	180
QY	172	PPEGIPPAEVE	231
DB	181	PPEGIPPAEVE	240
QY	232	SASAAVIVYNG	291
DB	241	STTATVIVYNG	299
QY	292	SSLLSVDSGSP	351
DB	300	CTTICPDVGAM	359
QY	352	HS-----	401
DB	360	QSEPVPVLEAS	417
QY	402	TSGFOPVSKPA	458
DB	418	TGCFHPVNFKA	476
QY	459	LCGGHHTLHSS	490
DB	477	LPSSLKVVYSST	536
QY	491	LPRGTSNMTY	549
DB	537	LPRDPGSSVSG	595
QY	550	GCOTLLSPVSG	609
DB	596	GTQTVLSPSVT	655
QY	610	BAPSHLYYQLE	659
DB	656	ETLNTPCYQLE	715
QY	670	YCLHDTHDAL	729
DB	716	YCLEDETPAL	775
QY	730	QETPFPHWNGT	789
DB	776	QETPFPHWNGS	834
QY	790	AEILLALESEAG	847
DB	835	GSLDTCSPAGS	894
QY	848	SFFASKPSPTAM	899
DB	895	NYFATKASPTG	946
RESULT 15			
AAW78900			
ID	AAW78900 standard; protein; 943 AA.		
XX			
AC	AAW78900;		

XX	25-MAR-2003	(revised)	
DT	21-DEC-1998	(first entry)	
XX	Rat UNC-5 homologue	UNC5H-2.	
XX	UNC-5; UNC5H-2; rat; netrin receptor; cell migration; axon guidance; diagnosis; therapy.		
OS	Rattus sp.		
XX	Key	Location/Qualifiers	
FT	Peptide	148..161	
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
FT	Misc-difference	753	
FT	/note= "encoded by CG"		
FT	Peptide	909..924	
FT	/note= "peptide used to raise rabbit polyclonal antisera"		
XX	W09837085-A1.		
XX	27-AUG-1998.		
XX	19-FEB-1998;	98WO-US003143.	
XX	19-FEB-1997;	97US-00808982.	
XX	(REGC ) UNIV CALIFORNIA.		
XX	Tessier-Lavigne M, Leonardo ED, Hinck L, Masu M, Keinomasu K;		
XX	WPI; 1998-495364/42.		
XX	N-PSDB; AAV52942.		
DR	Netrin-binding, vertebrate proteins - useful for diagnosis, therapy and the biopharmaceutical industry.		
XX	Claim 1; Page 24-26; 32pp; English.		
XX	UNC5H-1 and UNC5H-2 (see AAW78900) are rat homologues of Caenorhabditis elegans UNC-5 protein. Their amino acid sequences were deduced from an isolated unc5h cDNA clones (see AAV52940 and AAV52942) isolated from an B18 brain cDNA library. The predicted proteins show similarity with UNC-5, possess 2 predicted Ig-like domains and 2 predicted thrombospondin type-1 repeats, a predicted membrane spanning region, and a large intracellular domain. They are predicted to be involved in cell migration and axon guidance, and are characterised as receptor proteins for netrins. Human UNC5H-1 (see AAW78899) and UNC5H-2 (see AAW78901) proteins are also claimed. Vertebrate UNC-5 proteins may be produced recombinantly from transfected host cells. The invention also provides unc-5 hybridisation probes and primers, vertebrate UNC-5-specific binding agents such as specific antibodies, and methods of making and using the subject compositions in diagnosis (e.g. genetic hybridisation screens for vertebrate unc-5 transcripts), therapy (e.g. gene therapy to modulate vertebrate unc-5 gene expression) and in the biopharmaceutical industry (e.g. as immunogens, reagents for modulating cell guidance, reagents for screening chemical libraries for lead pharmacological agents, etc.).		
XX	(Updated on 25-MAR-2003 to correct PI field.)		
XX	Sequence 943 AA;		
Query Match 52.0%; Score 2490; DB 2; Length 943;			
Best Local Similarity 52.3%; Pred. No. 1.1e-206;			
Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;			
QY	9	PALLGIVLA	68
DB	21	PSLAGI-----	70
QY	69	PATQIPFKNGEW	128
DB	71	PATQIPFKNGEW	130





**This Page Blank (uspto)**

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	4545.5	95.0	898	2	US-08-808-982-5	Sequence 5, Appli
2	4545.5	95.0	898	3	US-09-308-902A-5	Sequence 5, Appli
3	2789.5	58.3	557	2	US-08-808-982-6	Sequence 6, Appli
4	2789.5	58.3	557	3	US-09-306-902A-6	Sequence 6, Appli
5	2490	52.0	943	2	US-08-808-982-7	Sequence 7, Appli
6	2490	52.0	943	3	US-09-306-902A-7	Sequence 7, Appli
7	294	6.1	102	2	US-08-808-982-8	Sequence 8, Appli
8	294	6.1	102	3	US-09-308-902A-8	Sequence 8, Appli
9	273.5	5.7	1172	1	US-08-313-288B-19	Sequence 19, Appl
10	249	5.2	441	3	US-08-985-526-3	Sequence 3, Appli
11	246	5.1	239	5	PT-0593-01652-1	Sequence 1, Appli
12	246	5.1	1170	1	US-08-313-288B-20	Sequence 20, Appl
13	235.5	4.9	218	3	US-08-985-526-1	Sequence 1, Appli
14	234	4.9	1395	3	US-09-540-245A-15	Sequence 15, Appl
15	216.5	4.5	1651	3	US-09-540-245A-18	Sequence 18, Appl
16	208.5	4.4	469	1	US-08-313-288B-15	Sequence 15, Appl
17	205.5	4.3	1381	3	US-08-540-245A-16	Sequence 16, Appl
18	198.5	4.1	788	2	US-08-918-914-4	Sequence 4, Appli
19	194.5	4.1	1069	4	US-09-877-730-2	Sequence 2, Appli
20	194.5	4.1	1150	4	US-09-877-730-8	Sequence 8, Appli
21	189	3.9	904	4	US-09-877-730-6	Sequence 6, Appli
22	189	3.9	985	4	US-09-877-730-10	Sequence 10, Appl
23	188	3.9	1297	3	US-09-540-245A-17	Sequence 17, Appl
24	186	3.9	380	4	US-09-877-730-4	Sequence 4, Appli
25	186	3.9	1266	4	US-08-506-236B-4	Sequence 4, Appli
26	168	3.5	481	4	US-09-130-491-8	Sequence 8, Appli
27	167.5	3.5	905	4	US-09-369-364A-9	Sequence 9, Appli





TELEFAX: (415) 343-4342  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 557 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-09-306-902A-6

Query Match 58.3%; Score 2789.5; DB 3; Length 557;  
 Best Local Similarity 96.4%; Pred. No. 2.7e-256;  
 Matches 538; Conservative 2; Mismatches 15; Indels 3; Gaps 3;  
 QY 344 NCTSLCVHSGAGPDVAVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTS 403  
 DB 1 NCTSLXVHTASGPDVAVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVADSSILTS 60  
 QY 404 GFQPVSIKSKADNPHLLTIQDPLS-TTITYGSLCPRQDGSPKFLTNHLLSPLGGG 462  
 DB 61 GFQPVSIKSKADNPHLLTIQDPLS-TTITYGSLCPRQDGSPKFLTNHLLSPLGGG 120  
 QY 463 RHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNNTYGTNFGRLMIPNTGISLLIP 522  
 DB 121 RHTLHSSPTSEAEFVSRLSTQNYFRSLPRGTSNNTYGTNFGRLMIPNTGISLLIP 180  
 QY 523 PDPAIRGKIYEIYTLHKPEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHCGEP 582  
 DB 181 PDPAIRGKIYEIYTLHKPEDVRLPLAGCQTLLSPVSCGPPGVLLTRPVILAMDHCGEP 240  
 QY 583 SPDSWSLRUKKSCGEGSWEDVHLGEEAPSHLYCYCOLEASACIYVTEOLGRFALVGEAL 642  
 DB 241 SPDSWSLRUKKSCGEGSWEDVHLGEEAPSHLYCYCOLEASACIYVTEOLGRFALVGEAL 299  
 QY 643 SVAAAKRLKLLFAPVACTSLENIYRVYCLDHTDALKKEVWOLEKOLGGQLIQEPRVLHF 702  
 DB 300 SVAAAKRLKLLFAPVACTSLENIYRVYCLDHTDALKKEVWOLEKOLGGQLIQEPRVLHF 359  
 QY 703 XDSYHNLRLSIHDVSPSLKSLVSYQIPIFYHWNQIQRVYLHCTFTTLERSVPSSTDLA 762  
 DB 360 XDSYHNLRLSIHDVSPSLKSLVSYQIPIFYHWNQIQRVYLHCTFTTLERSVPSSTDLA 419  
 QY 763 CKLWVQVGGGQSFNINIKDTRFAELLALAEAGVPALVGPFAKIPFLIROKILS 822  
 DB 420 CKLWVQVGGGQSFNINIKDTRFAELLALAEAGVPALVGPFAKIPFLIROKILS 479  
 QY 823 SLDPCCRGADWRTLAQKLHLDHSLGFFASKPSPTAMILNLWEARHFFPNGNLSQLAAVA 882  
 DB 480 SLDPCCRGADWRTLAQKLHLDHSLGFFASKPSPTAMILNLWEARHFFPNGNLSQLAAVA 539  
 QY 883 GLGQPDAGLFT-VSEAE 899  
 DB 540 GTXPAGRWLLSQCSAEAC 557

## RESULT 5

US-08-982-7  
 Sequence 7, Application US/08808982  
 Patent No. 5939271  
 GENERAL INFORMATION:  
 APPLICANT: Tessier-Lavigne, Marc  
 APPLICANT: Leonardo, E. David  
 APPLICANT: Hink, Lindsay  
 APPLICANT: Masu, Masayuki  
 APPLICANT: Kazuko, Keino-Masu  
 TITLE OF INVENTION: Netrin Receptors  
 NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: SCIENCE & TECHNOLOGY LAW GROUP  
 STREET: 268 BUSH STREET, SUITE 3200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA

COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/808,982  
 FILING DATE:  
 CLASSIFICATION: 530  
 ATTORNEY/AGENT INFORMATION:  
 NAME: OSMAN, RICHARD A  
 REGISTRATION NUMBER: 36,627  
 REFERENCE/DOCKET NUMBER: UC96-217  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 343-4341  
 TELEFAX: (415) 343-4342  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 943 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: not relevant  
 TOPOLOGY: not relevant  
 MOLECULE TYPE: peptide  
 US-08-808-982-7

Query Match 52.0%; Score 2490; DB 2; Length 943;  
 Best Local Similarity 52.3%; Pred. No. 2e-227;  
 Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;

QY 9 PALLGIVLAAMLURSGAQQSATVANPVGANPDLLPHLPEPEDEVYIVKKNPVLVCKAV 68  
 DB 21 PSLAGI-----DSGAQ---GLPDSFFSAPAEQLPHLLEPEDAIVKKNPVELHCRAF 70  
 QY 69 PATQIFFKCNGBWVRQDVHVIERSITDGSSEPTMEVRINVSQQQVEKVGLEEYWCQVA 128  
 DB 71 PATQIVFKCNGBWVSKGHVTQESLDEATGLRIEVRQVEVRSQQVELEFGLDYWCQVA 130  
 QY 129 WSSSGTTKSOKAYIRIARLRKNFEPLEAKESLEQIVLPCRPPGIGIPAEVEMURNED 188  
 DB 131 WSSSGTTKSRRAYIRIAYLRKNFDOEPLAKEVPLDHEVLLQCRPPGVPVAEVEMLKNE 190  
 QY 189 LVDPSSLDPNVYITRSHLVVRQARLADTANYTCVAKNIVARRSASAAVIVYVNGWSTW 248  
 DB 191 VIDPAQDTNFLTIDHNLIIRQARLSDTANYTCVAKNIVAKRSTTATVIVYVNGWSSW 250  
 QY 249 TMSVCSASCGRGWQVRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSDGSMSPSKW 308  
 DB 251 AEWSPCSNRCGRGWQKTRTCTNPAPLNGGAFCEGO-ACQKTACTTVCVPDVGAWTEWSK 309  
 QY 309 SACGLDCTHRSRECSDPAPRNGGECQGTDLDTNCTSDLCV---HSASGPE----- 358  
 DB 310 SACSTECALHRSRECMAPPNGGRDCSGTLLDSKNCTDGLCVLNORTLNDPKSRPLEPS 369  
 QY 359 -DVALYVGL-IAVAVCLVLLVILVYCRKKEGLSDVADSS-ILTSGFQPVSIKPSKA 415  
 DB 370 GDVALYAGLVAVFVVLAVLMAVGIVYRNCRDFDTIDTSSAALTGGPHVNFKTARP 429  
 QY 416 DNPHELL-TIQPDLSTTT-TYQGSCLPRQDGSPKFLTNHLLSPL----- 459  
 DB 430 SNPQLLHPSAPDLTASAGIYRGPVVALQDS-ADKIPMTNSPLDPLFSUKIVYOSSTI 488  
 QY 460 --GGG-----RHTLHSSPTSEAEFVSRLSTQNYFRSLPRG 494  
 DB 489 GSGAGLADGADLLGVLPGGTYPGDFSRDTHFLHRS-----ASLGSG-HLLGLPRD 538  
 QY 495 TSNMTYGTNFGRLMIPNTGISLLIPDAIPRGKIYEIYTLHKPEDVRLPLA-CQOT 553  
 DB 539 PSSSVSGTFCGLGRLTIPGTGVSLVLPNGAIPQGRFYDLYLRINKTEST-LPLSEGSQT 597  
 QY 554 LLSPIVSCGPPGVLLTRPVILAMDHCGEPDPSLSRLKKQSCGSEWQVHLGEGAPS 613

Db 598 VLSPSVTCGPTGLLLCRPVVLTVPCHCAEVIAGDWIFOLKTOAQHGWHE-EVVTLDDEETLN 656  
QY 614 HLYYCOLEASACVVFTEQGRFALVEALSVAAAKELKLLLPAPVACTSLEYNIRVYCLH 673  
Db 657 TPCYCOLEAKSCHILLDQGTFTFTGESYSRGAVRQLQAIAPALACTSLEYSLRVYCLE 716  
QY 674 DTHDALKEVVOLEKQGGOLIQEPRVLFHFKDSYHNLRLSLSDHVPSSLSKSLVSYQIEP 733  
Db 717 DTPAALKEVLELERTLGGVLEPEKTLFLFKDSYHNLRLSLSDHVPSSLSKSLVSYQIEP 776  
QY 734 FYHIWNGTORYLHCTFTLERHSLASTFTCKVCRQVEGEGQIFOLHTTLA-ETPAGSLD 793  
Db 777 FYHVWNGSQALHCTFTLERHSLASTFTCKVCRQVEGEGQIFOLHTTLA-ETPAGSLD 835  
QY 794 ALESEAGVPAL--VGPSAFKIPFLIROKLISSIDPPCRGADWRTLAOKLHLDLSHLSPPA 851  
Db 836 ALCSAPGNAATTQGLPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAOKLSMDRYLVNPA 895  
QY 852 SKPSTAMILNLWEARHFPNGNLSQLAAVAGLGOPDAGLFTVSEAE 899  
Db 896 TKASPTGVILDLWEARQDDGLNSLASALEEMGKSEMLVAMTTDGC 943

RESULT 6

US-09-306-902A-7  
; Sequence 7, Application US/09306902A  
; Patent No. 6277585

GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsay  
; Masu, Masayuki

TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 9

CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/306,902A  
; FILING DATE: 07-May-1999  
; CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217

TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342

INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 943 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: not relevant

MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 7:

US-09-306-902A-7  
Query Match 52.0%; Score 2490; DB 3; Length 943;  
Best Local Similarity 52.3%; Pred. No. 2e-227;  
Matches 496; Conservative 144; Mismatches 226; Indels 82; Gaps 19;

QY 9 PALGIVLAALRGGAQGSATVANPVGANPDLLPHFLVEPVDVYIVKPNVLVCKAV 68

Db 21 PSLAGI-----DSGAQ---GLPDSFPAPAEQPHFLLEPEDAYIVKKNFVELHCRAF 70  
QY 69 PATQIFFKCNQWVRQVHDVIERSTDGSGEPTMEVRINVSRQVQVBEKVFGLBEYWCQOVA 128  
Db 71 PATQIFYKCNQWVRQVGHVTOESLDEATGLRREVQIEVSQVQVBELEFGLDYWCQOVA 130  
QY 129 WSSSGTTTSOKAYIRIARLRKRNPEQPIAKEVSLBQGIIVLPCRPPEGIPPAEVEWLNED 188  
Db 131 WSSSGTTTSRRAYIRIAYLRKNDQEPLEAKEVPLDQEVLLQCRPPGVPVAVEWLNED 190  
QY 189 LVDPSLDPNVYITRHSILVVRQARLADTANYTCVAKNIVARRBSASAIVIVVNGWSWTW 248  
Db 191 VIDPAQDTNFIILIDHNLIIROARLSDTANYTCVAKNIVAKRRSTTATVIVVNGWSWSW 250  
QY 249 TEWSVCSACRGWOKRSCTNPAPLNGGAPCEGQVHDRTVSSILLVSDGSGWSPWSKW 308  
Db 251 AEWSPCSNRCRGWQKTRTCTNPAPLNGGAFCEGQ-ACQKTACTTCTVPCVDGANTWESKW 309  
QY 309 SACGLDCTHWSRECSDDPAPRNGECCGTDLDRNCTSDLCV---HSAASGE----- 358  
Db 310 SACSTECARWSRECMAPPQNGGRDCSGTLLDSXNCTDGLCVLNORTLNDPKSRPLEPS 369  
QY 359 -DVALYVGL-IAYAVCLVLLVLLVYCRKKEGLSDVADSS-ILTSGFQPVSIKPSKA 415  
Db 370 GDVALYAGLVAVFVVLAVLMAVGIVVYRRCRDFDTIDTSSAALTGGFHFVNFKTARP 429  
QY 416 DNPILL--TIQPDLSITT--TYGSLCPRQDQSPKQFQLTNGHLLSPL----- 459  
Db 430 SNPQLLHPSAPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLDPLSLKIKVYDSSTI 488  
QY 460 -GGG-----RHTLHSSPTSEAEFEVSRSLSTQNYRSLPRG 494  
Db 489 GSGAGLADGADLGLVLPPTYPGDFSRDTHFLHLS-----ASLGSQ-HLLGLPRD 538  
QY 495 TSNMTYGTFFNLGGRLMIPNTGISILIPDPAIPRGKIYEIYLTHKPDVRLPLA-GCQT 553  
Db 539 PSSVSVGTFGCLGRLTIPGTGVSLLVNGAIPQCKFYDLYLRINKTEST-LPLSEGSQT 597  
QY 554 LLSPIVSCGPPGVLTIRPVILAMDHCGEPPSPDSWSRLKQSCGSGWQDVVHLHGEAPS 613  
Db 598 VLSPSVTCGPTGLLLCRPVVLTVPCHCAEVIAGDWIFOLKTOAQHGWHE-EVVTLDDEETLN 656  
QY 614 HLYYCOLEASACVVFTEQGRFALVEALSVAAAKELKLLLPAPVACTSLEYNIRVYCLH 673  
Db 657 TPCYCOLEAKSCHILLDQGTFTFTGESYSRGAVRQLQAIAPALACTSLEYSLRVYCLE 716  
QY 674 DTHDALKEVVOLEKQGGOLIQEPRVLFHFKDSYHNLRLSLSDHVPSSLSKSLVSYQIEP 733  
Db 717 DTPAALKEVLELERTLGGVLEPEKTLFLFKDSYHNLRLSLSDHVPSSLSKSLVSYQIEP 776  
QY 734 FYHIWNGTORYLHCTFTLERHSLASTFTCKVCRQVEGEGQIFOLHTTLA-ETPAGSLD 793  
Db 777 FYHVWNGSQALHCTFTLERHSLASTFTCKVCRQVEGEGQIFOLHTTLA-ETPAGSLD 835  
QY 794 ALESEAGVPAL--VGPSAFKIPFLIROKLISSIDPPCRGADWRTLAOKLHLDLSHLSPPA 851  
Db 836 ALCSAPGNAATTQGLPYAFKIPLSIRQKICNSLDAPNSRGNDWRLLAOKLSMDRYLVNPA 895  
QY 852 SKPSTAMILNLWEARHFPNGNLSQLAAVAGLGOPDAGLFTVSEAE 899  
Db 896 TKASPTGVILDLWEARQDDGLNSLASALEEMGKSEMLVAMTTDGC 943

RESULT 7

US-08-808-982-8  
; Sequence 8, Application US/08808982  
; Patent No. 5939271  
GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; APPLICANT: Hink, Lindsay  
; APPLICANT: Masu, Masayuki

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/808,982  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: OSMAN, RICHARD A  
REGISTRATION NUMBER: 36,627  
REFERENCE/DOCKET NUMBER: UC96-217  
TELEPHONE: (415) 343-4341  
TELEFAX: (415) 343-4342  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 102 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-09-306-902A-8

Query Match 6.1%; Score 294; DB 2; Length 102;  
Best Local Similarity 56.4%; Pred. No. 4.1e-20;  
Matches 57; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

Qy 609 EAPSHLYYQLEASACVFTQGLGRFALVGEALSVAARKRLKLLFPAPVACTSLEYNIR 668  
Db 2 EETLNTPCYXQLEPRACKILLDQLGTGYVTGTSYRSVAKRLQAVFAPALCTSLVSLR 61

Qy 669 VYCLHDTHDALKEVQLEKQGLIQEPRVLHFKDSYHNL 709  
Db 62 VYCLEDTPVALKEVLEERTLGGYLVEEPKPLMFKDSYHNL 102

RESULT 8  
US-09-306-902A-8  
Sequence 8, Application US/09306902A  
Patent No. 6277585  
GENERAL INFORMATION:  
APPLICANT: Tessier-Lavigne, Marc  
Leonardo, E. David  
Hink, Lindsay  
Masu, Masayuki  
Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SEQUENCES: 8  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
STREET: 268 BUSH STREET, SUITE 3200  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:

APPLICANT: Kazuko, Keino-Masu  
TITLE OF INVENTION: Netrin Receptors  
NUMBER OF SE



US-08-313-288B-19

Query Match	5.7%;	Score 273.5;	DB 1;	Length 1172;
Best Local Similarity	28.9%;	Fred. No. 2.1e-16;		
Matches	74;	Conservative	27;	Mismatches 110; Indels 45; Gaps 8;
QY	209	QOARLADTANYTCVAKNIVARRRSASAA-VIVYVNGGWSTWTEMSVGSASCGRGQKRSR	267	
DB	403	QGRSCDVNTNTCLGPSIQTRACSLSKCDTRIRQDGGKSHWSPWSSCSVTGCVGNITRIR	462	
QY	268	SCTNPAFLNGGACGEQNVHDIRTVSSLLVSDGWSFWKWSACGLDCT---HWRSRECS	324	
DB	463	LCNSPVPQMGKKNGSGRRTKACQCAPCPIDGWSFWSPWSPSACTVTCAGGIRERTVCN	522	
QY	325	DPAPRNGGECQGTDLTRNCTSLCVHWSASGPDVAIVYGLIAVAVCLVLLLLVLIVY	384	
DB	523	SPEFQYGGKACVGDVQBRQMCNKRSC-----PVDGCLSNCFPGAQC-----	564	
QY	385	CRKKEGLDSVADSSILITSGFQPSVI--KPSKADNPHELLIQDPLSTTT-----TYQ	434	
DB	565	-----SSFPGDS-WSCGFCFVGFTLNGTHCEDLDECALVPDQICFSTKVPKPCVNTQP	615	
QY	435	GSLC-----PRQDGPSP	446	
DB	616	GFHCLPCPPRYRGNOP	631	

## RESULT 10

US-08-985-526-3

Sequence 3, Application US/08985526

Patent No. 6080728

GENERAL INFORMATION:

APPLICANT: Mixson, James A

TITLE OF INVENTION: CARRIER-DNA COMPLEXES CONTAINING DNA

TITLE OF INVENTION: ENCODING ANTI-ANGIOGENIC PEPTIDES AND THEIR USE IN GENE

TITLE OF INVENTION: THERAPY

NUMBER OF SEQUENCES: 43

CORRESPONDENCE ADDRESS:

ADDRESSEE: Connolly, Bove, Lodge, & Hutz

STREET: 1220 Market Street, P.O. Box 2207

CITY: Wilmington

STATE: Delaware

COUNTRY: U.S.A.

ZIP: 19899

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/985,526

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/608,845

FILING DATE: 16-JUL-1996

ATTORNEY/AGENT INFORMATION:

NAME: McMorow Jr., Robert G

TELECOMMUNICATION INFORMATION:

TELEPHONE: (302) 658-9141

TELEFAX: (302) 658-5613

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 441 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-985-526-3

```

132 FKQDGGSHWSPWSSCVTCGDGVITRITLQNSPSPQMNQKPCGEARETKACKDACP 191
114 EKVGLEEWQOCVAVSSSGTTKSKAVIRIARLRYNFEQELAKEVSLQGIIVLPCRPP 173
192 NGMGPSWPDWICSVTCGGGVQKRSRLCV--DSMTENKELANELR-----RPP 239
174 ---EGIPPAEVEUARNED-LVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVA 228
240 LCYHNG-----VOYRNNEEWIVDSCTE-----CHCQNSVT 269
229 RRRSASAAVIVYNG-----GWSTWTWVSVCSASCGRGWQKESRSC 269
270 ICKVSCFIMPICSNATVPDGECCPRCWPDSADDCGWSPWSEWTSCTSCGNGIQKRGSC 329
270 TNPAPLNGAFCEGONVHDTV-----SSLLYSVQGSWSPWKSWSACGLDC---THWRSRE 322
330 DS----LNNR--CEGSSVUTRCHIQECDKRFQDGGWSHWSPWSSCSVTGCGDVITRITL 384
323 CSDPAPRNGBECOGTDLDRNCTSDLC-VHSASGP 357
385 CNSPSPQWNGKPCGEARETKACKDACPINGWGP 420

RESULT 11
PCT-US93-01652-1
; Sequence 1, Application PC/TUS9301652
; GENERAL INFORMATION:
; APPLICANT: Bouck, Noel P.
; APPLICANT: Polverini, Peter J.
; APPLICANT: Good, Deborah J.
; APPLICANT: Frazier, William A.
; TITLE OF INVENTION: Method and Composition for
; TITLE OF INVENTION: Inhibiting Angiogenesis
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Iilton, Fallon, Lungmus & Chestnut
; STREET: 100 South Wacker Drive, Suite 960
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/01652
; FILING DATE: 19930222
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/841,656
; FILING DATE: 24-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/464,369
; FILING DATE: 12-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fentress, Susan B.
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92005-PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312)-456-8000
; TELEFAX: (312)-456-7776
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 239 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
PCT-US93-01652-1

```



1 ; CURRENT FILING DATE: 2000-03-31  
2 ; PRIOR APPLICATION NUMBER: 60/065,544  
3 ; PRIOR FILING DATE: 1997-11-14  
4 ; PRIOR APPLICATION NUMBER: 60/081,057  
5 ; PRIOR FILING DATE: 1998-04-07  
6 ; NUMBER OF SEQ ID NOS: 20  
7 ; SOFTWARE: Patent In Ver. 2.0  
8 ; SEQ ID NO 15  
9 ; LENGTH: 1395  
10 ; TYPE: PRT  
11 ; ORGANISM: Drosophila melanogaster  
12 US-09-540-245A-15

Query Match 4.9%; Score 234; DB 3; Length 1395;  
Best Local Similarity 20.5%; Pred. No. 1.6e-12;  
Matches 188; Conservative 103; Mismatches 262; Indels 364; Gaps 46;

QY 4 RPLGPAALIGIVLAALRGGAQASATVANPFGA-----NPDLLPHFLVPEDEYVIVKN 58  
DB 28 RMWLLPAWLLLVLA-----SNGLPVAVRGVQSPRIIEH---PTDLVVKKN 70

QY 59 KPVLIVCK--AVPATQIFPKNGEVRVQDVHVIERSSTGSGSEPTMEVRINVS- 110  
DB 71 EPATLNCKVEGPEPTI-----EWFK-----DGEF---VSTNEKKSHRVQF 108

QY 111 -----QVEKVFGLBEEYKCOQVAMSSSGTTKSKQAVIRIARLRKNFEQELAKEV 160  
DB 109 KDCALFFRTMCKEKGQDGEY--CVAKNRVQAVSHASQIAVLRRDFFRPRDTRV 166

QY 161 SLEQGIPLPCRPPEGIPPAEVEWLRN-----EDLVDPSL--DENVTITREHSLVVRQARLA 214  
DB 167 AKGETALLECQPKGIPETLWIKDGVPLDLKAMSGASSRVRIVDGNLLISNVPEI 226

QY 215 DTANTYCVAKNIVARRRSASAIVIVYN-----GWSSTWTEWSVCASGCR 260  
DB 227 DEGNKYKIAQNLVGTRESSYAKLIQVKEPYFMKEPRQDQVMLYQOTATF-----HCSV 278

QY 261 G-----WQK-----RSRCTNPAPLNGGAF-CEGON----- 285  
DB 279 GGDPPPKVLWKKEBNI PVSRRARILHDEKSLISNITPTDEGYTCEAHNNVGOISARAS 338

QY 286 --VH-----DRTVSSLLVSDG-----WSWP-----WSKWSACGL--DCTHWSRE 322  
DB 339 LIVHAPPNTKPSNKKVGLNGVQVQLPCWASGNPPSPVETKEGVSTLMPFNSSHGROV 398

QY 323 CSDPAPRNGEGECQGTDLTRNCTSDLCVHSAGSEPDVALYVGLIAVAVLLVLL 382  
DB 399 AADGT-----LQITDV-----RQDEGY--VCSAF----- 422

QY 383 VYCRKKEGLSDVADSSILTSFGQPVSIKPSKADN--PHLLTIOPDLSTTTTYOGLS--- 437  
DB 423 -----SVDSSTVR-----VFLQVSSVDERPPPIIQGP--ANQTLPGKSVATL 464

QY 438 -CFRQDGPSPKQF-LTNGHLLPLGGRHLLHSSPTSEAEFVSRSLSTONYFRSLPRGT 495  
DB 465 PCRAITGNPSRIKWFHDGHAQ--AGNRYSIIQSSLRVDDQLQSDSGYITCTASGERGE 522

QY 496 SNNYTGTFNGLGRMLMPTNGISLLIPDPAIPRGKIYEIYLHLKPEDVRLPLAGCQTL 555  
DB 523 TS-----WAATLIVKEPKGSTSLHRAA----- 543

QY 556 SPVIVSGPPGVLLTRPVILAMDHCGEFPSPDSLSLR-LKKQSCGS----- 599  
DB 544 DPSTVPAPPGT-----PKVLNV-----SRTSISLRWAKSQEPKPGAVGPIIYTVVEFSPD 593

QY 600 -----WEQDVLHLGEEAPSHLYYQLEASACYVF-----TEOLGR 634  
DB 594 LOTGMIVAAHRVGD---TQVTTISGLTPGTSYVFLVRAENTQGISVPSGLSNVIKTIADP 650

QY 635 FALVGEALSA--AAKRLKLLFAPVACTSLEYNIRVYCLHDTDALKEVVVQLEKQLG 690  
DB 651 DAASANDLSAARTLLTGKVELIDASAINASAVRLE---WMLHVSAD-----EKYVE 699

QY 691 GQLIQEPRVLHFKDS-----YHNRL-----SIHVD 716  
DB 700 GLRL-----HYKDSVPSAQYHSITVMDASAESFVGNLKKYKYEPFLTPFFETIEQ 753

QY 717 PSSLWKSLLVSYQEP 733  
DB 754 PSN---SKTALTYEDVP 767

RESULT 15  
US-09-540-245A-18  
; Sequence 18, Application US/09540245A  
; Patent No. 6270984  
; GENERAL INFORMATION:  
; APPLICANT: Goodman, Corey  
; APPLICANT: Kid, Thomas  
; APPLICANT: Brose, Katja  
; APPLICANT: Tessier-Lavigne, Marc  
; TITLE OF INVENTION: Modulating Robo: Ligand Interactions  
; FILE REFERENCE: B98-031-3  
; CURRENT APPLICATION NUMBER: US/09/540,245A  
; CURRENT FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 60/065,544  
; PRIOR FILING DATE: 1997-11-14  
; PRIOR APPLICATION NUMBER: 60/081,057  
; PRIOR FILING DATE: 1998-04-07  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 18  
; LENGTH: 1651  
; TYPE: PRT  
; ORGANISM: human  
; US-09-540-245A-18

Query Match 4.5%; Score 216.5; DB 3; Length 1651;  
Best Local Similarity 20.7%; Pred. No. 1e-10;  
Matches 213; Conservative 139; Mismatches 381; Indels 295; Gaps 54;

QY 41 DLLPHFLVEPEYIVKPKVLLVCKAV--PATQIFKNGEVRVQDVHVIERSSTGSG 98  
DB 65 DFPFRIVEHPSDLIYSKGEPAATLNCKAEGRPPTI-----EWYKGEVERETKDDPRSH 118

QY 99 EPTME-----VRINVSQOQVEKVFGLBEEYKCOQVAMSSSGTTKSKQAVIRIARLRKNFE 152  
DB 119 RMLPSGSLFFLIRIVHGRKSRP-----DEGVTVCVARNYLGEAVSHNASLEVAILLRDDR 173

QY 153 QBFLAKESLEQGIPLPCRPPEGIPPAEVEWLRNEDLVDPSLDPNVYIT-REHSLVVRQA 211  
DB 174 QNPSDVMVAVGPAVMEQPPRGHPEPTISWKKGSPILD---DKDERITIRGGKLMITVT 230

QY 212 RLADTANTYCVAKNIVARRRSASAIVIVYNGW-----STWTEWSV---CSASCGRG 261  
DB 231 RKSDAGKVCVGTNVMGERESEVABLTVLERFSVKRPSNLAVTVDDSAEFKCEA---RG 287

QY 262 -----WQK-----RSRCTNPAPLNGGAFCEGONVHRTVSSLLV 296  
DB 288 DPVPTVRWKDGEPLPKSEYIRDHDTLKIRKVT--AGDMGSYTCVAENMVKAESATL 345

QY 297 SYDGS-----WSPWSKWSACGLDCTHWSRECSDDPAP-----RNGGEEC-----QGTDL 340  
DB 346 TVQEPFPHFVKPRDQVVALGRVT--FQCEATGNPQPAIFWRREGSQNLLFSYQPPQSSSR 404

QY 341 DTRNCTSDLCVHSAGSEPDVALYVGLIAVAVCLVLLVLLVYCRKKEGLSDVADSSI 400  
DB 405 FSVSQTGLTITNVQ-RSDVGYY-----ICQTL-----NVA-GSI 437

QY 401 LTSFGQPVSIKSKADNPHLLTIOPDLSTTTTYOGLS---CFRQDGPSPKF----- 448  
DB 438 ITKAY--LEVTVIADRPVPPVIRQGPVNTQVAVDGTFLVSCVATGSPVTILWRKDGVLV 495

QY 449 -----QLTNGHL---LSPLGG-GRHTLHSSPTSEAEFVSRSLSTONYFRSLPRGTS 496  
DB 496 STQDSRIKOLENGVLQIRYAKLGDITGRYTCIASTPSGEA----- 534



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:29:07 ; Search time 66 Seconds  
(without alignments)  
4383.295 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLWPALLGIVLAWL.....AVAGLGQDPAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 32179191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/2/pubaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4787	100.0	899	11	US-09-970-944-2
2	4698.5	98.2	898	12	US-09-918-779-2
3	4698.5	98.2	898	16	US-10-624-932-2
4	4545.5	95.0	898	10	US-09-933-261-5
5	4545.5	95.0	898	11	US-09-970-944-13
6	4545.5	95.0	898	14	US-10-256-702-5
7	4545.5	95.0	898	14	US-10-240-154-16
8	4379.5	91.5	842	16	US-10-311-623-1
9	2819	58.3	544	11	US-09-970-944-14
10	2789.5	58.3	557	10	US-09-933-261-6
11	2789.5	58.3	557	14	US-10-256-702-6
12	2705.5	56.5	931	11	US-09-970-944-15
13	2705.5	56.5	931	12	US-10-087-684-35
14	2705.5	56.5	931	12	US-09-972-211-121
15	2705.5	56.5	931	12	US-10-037-417-117

16	2705.5	56.5	931	12	US-10-096-625-121
17	2705.5	56.5	1010	12	US-10-218-779-35
18	2680.5	56.0	931	11	US-09-970-944-16
19	2680.5	56.0	931	12	US-09-972-211-125
20	2680.5	56.0	931	12	US-10-096-625-125
21	2677.5	55.9	931	12	US-10-037-417-120
22	2673.5	55.8	931	11	US-09-970-944-17
23	2673.5	55.8	931	12	US-10-087-684-36
24	2673.5	55.8	931	12	US-10-218-779-36
25	2673.5	55.8	931	12	US-09-972-211-122
26	2673.5	55.8	931	12	US-10-037-417-118
27	2673.5	55.8	931	12	US-10-037-417-119
28	2673.5	55.8	931	12	US-10-096-625-122
29	2497	52.2	945	12	US-10-087-684-34
30	2497	52.2	945	12	US-10-218-779-34
31	2497	52.2	945	12	US-09-972-211-124
32	2497	52.2	945	12	US-10-037-417-121
33	2497	52.2	945	12	US-10-096-625-124
34	2493	52.1	945	12	US-10-087-684-33
35	2493	52.1	945	12	US-10-218-779-33
36	2493	52.1	945	12	US-09-972-211-123
37	2493	52.1	945	12	US-10-096-625-123
38	2490.5	52.0	946	12	US-09-972-211-56
39	2490.5	52.0	946	12	US-10-096-625-56
40	2490	52.0	943	10	US-09-933-261-7
41	2490	52.0	943	14	US-10-256-702-7
42	2484	51.9	933	12	US-10-087-684-2
43	2484	51.9	933	12	US-10-087-684-4
44	2484	51.9	933	12	US-10-218-779-2
45	2484	51.9	933	12	US-10-218-779-4

ALIGNMENTS

RESULT 1  
US-09-970-944-2  
; Sequence 2, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herrman, John L  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an  
; TITLE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; PRIOR FILING DATE: 2002-05-02  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 899  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-970-944-2

Query Match	100.0%	Score 4787;	DB 11;	Length 899;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 899;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAVRPGLWPALLGIVLAWL	GSQAQSATVANVPGANPDLLPHFLVPEPDVYIVKNKP	60
Db	1	MAVRPGLWPALLGIVLAWL	GSQAQSATVANVPGANPDLLPHFLVPEPDVYIVKNKP	60
QY	61	VLLVCKAVPATQIEFFKNGEWMVROVDHVEBSTDSSGSEPTMEVRIINVSROOVKEVFGLE	120	
Db	61	VLLVCKAVPATQIEFFKNGEWMVROVDHVEBSTDSSGSEPTMEVRIINVSROOVKEVFGLE	120	
QY	121	EYWCQCVAWSSGTTKQKAVIRIARLRKNFQEQPLAKEVLSLEQIGVLPFCPPPGIPPAE	180	
Db	121	EYWCQCVAWSSGTTKQKAVIRIARLRKNFQEQPLAKEVLSLEQIGVLPFCPPPGIPPAE	180	

QY 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240  
Db 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240  
QY 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300  
Db 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300  
QY 301 SWSPMKWSACGLDCTHWRSCSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360  
Db 301 SWSPMKWSACGLDCTHWRSCSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360  
QY 361 ALYVGLIAVAVCLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420  
Db 361 ALYVGLIAVAVCLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420  
QY 421 LTIQDLSLTITTYQGSCLPRQDGPSPKQFTNGHLLSPGGGRHTLHSSPTSSEAEFVS 480  
Db 421 LTIQDLSLTITTYQGSCLPRQDGPSPKQFTNGHLLSPGGGRHTLHSSPTSSEAEFVS 480  
QY 481 RLSTQNYFRSLPRGTSNNYTGTFNFGRLMIPNTGISILLIPDPAIPRGKIYEIYLTJHK 540  
Db 481 RLSTQNYFRSLPRGTSNNYTGTFNFGRLMIPNTGISILLIPDPAIPRGKIYEIYLTJHK 540  
QY 541 PEDVRLPLAGCOTLLSPVSCGPGVLTTRPVILAMDHCGEPPDSWSILRKKQCEGSM 600  
Db 541 PEDVRLPLAGCOTLLSPVSCGPGVLTTRPVILAMDHCGEPPDSWSILRKKQCEGSM 600  
QY 601 EQDVLHLGEEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAAXRLKLLFAPVAC 660  
Db 601 EQDVLHLGEEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAAXRLKLLFAPVAC 660  
QY 661 TSLEYNIRVYCLHDTKALKEVVQLEKQGGQLIQEPVRLHPKDSYHNLRLSIHDPVSSL 720  
Db 661 TSLEYNIRVYCLHDTKALKEVVQLEKQGGQLIQEPVRLHPKDSYHNLRLSIHDPVSSL 720  
QY 721 WSKLLVSYQETPFYHNGTORYLHCTFTLERSVPSTSDACKLWVQVGDGOSFSIN 780  
Db 721 WSKLLVSYQETPFYHNGTORYLHCTFTLERSVPSTSDACKLWVQVGDGOSFSIN 780  
QY 781 FNIKTDFABELLALLESAGVPALVGPFAKIPFLIRQKIISLPPCRRGADMTLQAK 840  
Db 781 FNIKTDFABELLALLESAGVPALVGPFAKIPFLIRQKIISLPPCRRGADMTLQAK 840  
QY 841 LHLDSLFFASKPSPTAMILNWEARHPFNGLSOLAAVAGLQPDAGLFTVSEAE 899  
Db 841 LHLDSLFFASKPSPTAMILNWEARHPFNGLSOLAAVAGLQPDAGLFTVSEAE 899

RESULT 2

US-09-918-779-2  
; Sequence 2, Application US/09918779  
; Publication No. US2003064369A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier, Raymond  
; APPLICANT: Padigar, Muralidhara  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Shimkets, Richard  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Li, Li  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Grosse, William  
; APPLICANT: Alsobrook, John  
; APPLICANT: Lepley, Denise  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Ellerman, Karen  
; APPLICANT: MacDougall, John  
; APPLICANT: Stone, David

; APPLICANT: Smithson, Glenda  
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-074 US  
; CURRENT APPLICATION NUMBER: US/09/918,779  
; CURRENT FILING DATE: 2001-07-30  
; PRIOR APPLICATION NUMBER: 60/221,409  
; PRIOR FILING DATE: 2000-07-28  
; PRIOR APPLICATION NUMBER: 60/222,840  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,752  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,762  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,770  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/223,769  
; PRIOR FILING DATE: 2000-08-08  
; PRIOR APPLICATION NUMBER: 60/225,146  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/225,392  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/225,470  
; PRIOR FILING DATE: 2000-08-15  
; PRIOR APPLICATION NUMBER: 60/225,697  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: 60/263,662  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: 60/281,645  
; PRIOR FILING DATE: 2001-04-05  
; NUMBER OF SEQ ID NOS: 61  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 898  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-918-779-2

Query Match 98.2%; Score 4698.5; DB 12; Length 898;  
Best Local Similarity 98.7%; Pred. No. 0;  
Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;  
QY 1 MAVRGLWPALLGIVLAALRGSGAQQSATVANPVPFANPDLLPHFLVEPEDVYIVKNKP 60  
Db 1 MAVRGLWPALLGIVLAALRGSGAQQSATVANPVPFANPDLLPHFLVEPEDVYIVKNKP 60  
QY 61 VLLVCKAVPATQIPKCNQWVRQVDHVIERSITDSSGSEPTMEVRINVSROQVEKVFGL 120  
Db 61 VLLVCKAVPATQIPKCNQWVRQVDHVIERSITDSSGSEPTMEVRINVSROQVEKVFGL 120  
QY 121 EYWCOCVAVSSGTTKSKAVIRIARLKNFEQEPFLAKEVSLQEGIVLPCRPPEGIPPAE 180  
Db 121 EYWCOCVAVSSGTTKSKAVIRIARLKNFEQEPFLAKEVSLQEGIVLPCRPPEGIPPAE 180  
QY 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240  
Db 181 VEWLNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240  
QY 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHRTVSSLLVSDG 300  
Db 241 VNGGWSTWTSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNV-OKTACATLCFVDG 299  
QY 301 SWSPMKWSACGLDCTHWRSCSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 360  
Db 301 SWSPMKWSACGLDCTHWRSCSDPAPNGGECQGTDLTRNCTSDLCVHSASGPDV 359  
QY 361 ALYVGLIAVAVCLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 420  
Db 361 ALYVGLIAVAVCLVLLVILVYCRKKEGLSDVADSSILTSFGQPSIKPSKADNPHL 419  
QY 421 LTIQDLSLTITTYQGSCLPRQDGPSPKQFTNGHLLSPGGGRHTLHSSPTSSEAEFV 479  
Db 421 LTIQDLSLTITTYQGSCLPRQDGPSPKQFTNGHLLSPGGGRHTLHSSPTSSEAEFV 479

```
QY 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGRLMIPNTGISLIPDPAIPRGKIYEIYLTGH 539
Db 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGRLMIPNTGISLIPDPAIPRGKIYEIYLTGH 539
QY 540 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGFSI 599
Db 540 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGFSI 599
QY 600 WEQDVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAAAKLKLFPAPVA 659
Db 600 WE-DVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAAAKLKLFPAPVA 658
QY 660 CTSLEYNIRVYCLDHTDHALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDVPSS 719
Db 659 CTSLEYNIRVYCLDHTDHALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDVPSS 718
QY 720 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDLACKLWVWQVEGDGQSF 779
Db 719 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDLACKLWVWQVEGDGQSF 778
QY 780 NFNITKDTTRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db 779 NFNITKDTTRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838
QY 840 KLHLDHSLSPFASKPSPMTAMLNILWEARHPNGNLISOLAAVAGLGQPDAGLFTVSEAC 899
Db 839 KLHLDHSLSPFASKPSPMTAMLNILWEARHPNGNLISOLAAVAGLGQPDAGLFTVSEAC 898

RESULT 3
US-10-624-932-2
; Sequence 2, Application US/10624932
; Publication No. US20040096877A1
; GENERAL INFORMATION:
; APPLICANT: Taupier, Raymond
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Rastelli, Luca
; APPLICANT: Spaderna, Steven
; APPLICANT: Shimkets, Richard
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Spytek, Kimberly
; APPLICANT: Shenoy, Suresh
; APPLICANT: Li, Li
; APPLICANT: Gusev, Vladimir
; APPLICANT: Grosse, William
; APPLICANT: Alsobrook, John
; APPLICANT: Lepley, Denise
; APPLICANT: Burgess, Catherine
; APPLICANT: Gerlach, Valerie
; APPLICANT: Ellerman, Karen
; APPLICANT: MacDougall, John
; APPLICANT: Stone, David
; APPLICANT: Smithson, Glenda
; TITLE OF INVENTION: Novel Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-074 US
; CURRENT APPLICATION NUMBER: US/10/624,932
; CURRENT FILING DATE: 2003-07-21
; PRIOR APPLICATION NUMBER: 09/918,779
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: 60/221,409
; PRIOR FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 60/222,840
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: 60/223,752
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,762
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,770
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/223,769
; PRIOR FILING DATE: 2000-08-08
; PRIOR APPLICATION NUMBER: 60/225,146
; PRIOR FILING DATE: 2000-08-14
```

```
; PRIOR APPLICATION NUMBER: 60/225,392
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: 60/225,470
; PRIOR FILING DATE: 2000-08-15
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 61
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 898
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-624-932-2

Query Match 98.2%; Score 4698.5; DB 16; Length 898;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 888; Conservative 2; Mismatches 7; Indels 3; Gaps 3;

QY 1 MAVRPGIWPALLGIVLAAWLGRSGAQQSATVANVPFGANPDLLPHFLVEPEDYIIVKNKP 60
Db 1 MAVRPGIWPALLGIVLAAWLGRSGAQQSATVANVPFGANPDLLPHFLVEPEDYIIVKNKP 60
QY 61 VLLVCKAVPATQIIPFKCNGEWVRQVHVIERSTDGSGSEPTMEVRINVSQQVEKVGLE 120
Db 61 VLLVCKAVPATQIIPFKCNGEWVRQVHVIERSTDGSGSEPTMEVRINVSQQVEKVGLE 120
QY 121 EYWCQCVAMSSSCTTKSQKAYIRIARLNKFEQELAKEVSLBQGIPLCRPPEGIPPAE 180
Db 121 EYWCQCVAMSSSCTTKSQKAYIRIARLNKFEQELAKEVSLBQGIPLCRPPEGIPPAE 180
QY 181 VEWLNRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
Db 181 VEWLNRNEDLVDPSLDPNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIV 240
QY 241 VNGGWSTWTSVCSASCGRWQKRSRSCNPNAPLNGGAFCEGQNVHDRTVSSLLVSDG 300
Db 241 VNGGWSTWTSVCSASCGRWQKRSRSCNPNAPLNGGAFCEGQNVHDRTVSSLLVSDG 300
QY 301 SWSPWKSWSACGLDCTHWRRECSDPAPRNGGECQGTDLDRNCTSDLCVHSASGEDV 360
Db 300 SWSPWKSWSACGLDCTHWRRECSDPAPRNGGECQGTDLDRNCTSDLCVHSASGEDV 359
QY 361 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVDADSSILTSGFQFVSKPADNPHL 420
Db 360 ALYVGLIAVAVCLVLLVLLVILVYCRKKEGLSDVDADSSILTSGFQFVSKPADNPHL 419
QY 421 LTIQPDLS-TTTTYQGSICPRQDGPSPKFQITNGHLLSLPLGGGHTLHHSPTSEAEFF 479
Db 420 LTIQPDLS-TTTTYQGSICPRQDGPSPKFQITNGHLLSLPLGGGHTLHHSPTSEAEFF 479
QY 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGRLMIPNTGISLIPDPAIPRGKIYEIYLTGH 539
Db 480 SRLSTQNYFRSLPRGTSNNMTYGTFFNLGGRLMIPNTGISLIPDPAIPRGKIYEIYLTGH 539
QY 540 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGFSI 599
Db 540 KPEDVRLPLAGCQTLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGFSI 599
QY 600 WEQDVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAAAKLKLFPAPVA 659
Db 600 WE-DVLHLGGEAPSHLYYCOLEASACVVFTEQGRFALVGEALSVAAAKLKLFPAPVA 658
QY 660 CTSLEYNIRVYCLDHTDHALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDVPSS 719
Db 659 CTSLEYNIRVYCLDHTDHALKEVVQLEKQGGOLIOBPRVLHFKDSVHNLRLSIHDVPSS 718
QY 720 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDLACKLWVWQVEGDGQSF 779
Db 719 LKSKLLVSYQEIIPFFYHINWGTQRYLHCTFTLBRVSPSTSDLACKLWVWQVEGDGQSF 778
QY 780 NFNITKDTTRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 839
Db 779 NFNITKDTTRFAELLALSEAGVPALVGPSPAFKIPFLIRQKIISSLDPPCRRGADWRTLAQ 838
```



QY 840 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAAVAGLQPDAGLFTVSEAC 899  
DB 839 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAAVAGLQPDAGLFTVSEAC 898

## RESULT 4

US-09-933-261-5  
; Sequence 5, Application US/09933261  
; Publication No. US20030040046A1  
; GENERAL INFORMATION:  
; APPLICANT: Tessier-Lavigne, Marc  
; Leonardo, E. David  
; Hink, Lindsay  
; Masu, Masayuki  
; Kazuko, Keino-Masu  
; TITLE OF INVENTION: Netrin Receptors  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
; STREET: 268 BUSH STREET, SUITE 3200  
; CITY: SAN FRANCISCO  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; FILING DATE: 20-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/808,982  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: OSMAN, RICHARD A  
; REGISTRATION NUMBER: 36,627  
; REFERENCE/DOCKET NUMBER: UC96-217  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 343-4341  
; TELEFAX: (415) 343-4342  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 898 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: No. US20030040046A1 Relevant  
; TOPOLOGY: No. US20030040046A1 Relevant  
; MOLECULE TYPE: peptide  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-933-261-5

Query Match 95.0%; Score 4545.5; DB 10; Length 898;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGIWPALLGIVLAALRGSGAQSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60  
DB 1 MAVRPGIWPVLLGIVLAALRGSGAQSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60  
QY 61 VLLVCKAVPATQIFKCNKGEWVRQVDHVIERTSTGSGSEPTMEVIRINVSQQVEKVFGL 120  
DB 61 VLLVCKAVPATQIFKCNKGEWVRQVDHVIERTSTGSGSEPTMEVIRINVSQQVEKVFGL 120  
QY 121 EYWCQCVAWSSGTTKQKAYIRIARLKNFPEGLAKEVSLQGVLPQCRPEGIPPAE 180  
DB 121 EYWCQCVAWSSGTTKQKAYIRIARLKNFPEGLAKEVSLQGVLPQCRPEGIPPAE 180  
QY 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSAAVIVY 240  
DB 181 VEWLRNEDLVDPSPDPNVYITREHSLVVRQARLADTANYTCVAKNIIVARRSAAVIVY 240

QY 241 VNGCWSTWTSVCSASCGRCQWQKRSRSCCTNPAPINGGAFCEGQNVHRTVTSLSLVSDG 300  
DB 241 VNGCWSTWTSVCSASCGRCQWQKRSRSCCTNPAPINGGAFCEGQNV-OKTACATLCPVDG 299  
QY 301 SWSPWSKWSACGLDCTHWSRRECSDDPAPRNGGEECCQGTDLDRNCTSDLCVHSASGPEDV 360  
DB 300 SWSWSKWSACGLDCTHWSRRECSDDPAPRNGGEECCGADLDRNCTSDLCUHTASCPEV 359  
QY 361 ALYVGLIAVAVCLVLLILVILVYCRKKEGLSDVADSSILTSFGPQPSVSKPSKADNPHL 420  
DB 360 ALYIGLVAVAVCLFILLALGLIYCRKKEGLSDVADSSILTSFGPQPSVSKPSKADNPHL 419  
QY 421 LTIQPDLS-TTTTQGSCLPRQDQSPKPFQLTNGHLLSPGLGGRHHTLHSSPTSAEBFV 479  
DB 420 LTIQPDLS-TTTTQGSCLPRQDQSPKPFQLTNGHLLSPGLGGRHHTLHSSPTSAEBFV 479  
QY 480 SRLSTQNYFRSLPRGTSMNTYGTNFGGLGRIMIPNTGISLLIPDAIPRGKIYEIYLTILH 539  
DB 480 SRLSTQNYFRSLPRGTSMNAYGTNFGGLGRIMIPNTGISLLIPDAIPRGKIYEIYLTILH 539  
QY 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDMSLRKKQSCBGS 599  
DB 540 KPEDVRLPLAGCQTLSPVSCGPPGVLLTRPVILAMDHCGEPPSDMSLRKKQSCBGS 599  
QY 600 WEQDVTLHGEAPSHLYYCOLEASACVVFTEOLGRFALVGEALSVAAAKRLKLLFAPVA 659  
DB 600 WE-DVTLHGEESPSHLYYCOLEAGACYVFTEQLGRFALVGEALSVAAATKRLKLLFAPVA 658  
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGLIQPRVLHFKDSYHNLRSLSHDVPSS 719  
DB 659 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGLIQPRVLHFKDSYHNLRSLSHDVPSS 718  
QY 720 LWSKLLVSYQEIPIFYHWNQTQYLRHCTFTLERINASTSLACKLWVWQVEGDQSPSI 779  
DB 719 LWSKLLVSYQEIPIFYHWNQTQYLRHCTFTLERINASTSLACKLWVWQVEGDQSPSI 778  
QY 780 NFNITKDTFAELLALAESEAGVPALVGFSAFKIFFLIROKIISSLDPPCRGADWRTIAQ 839  
DB 779 NFNITKDTFAELLALAESEAGVPALVGFSAFKIFFLIROKIISSLDPPCRGADWRTIAQ 838  
QY 840 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAAVAGLQPDAGLFTVSEAC 899  
DB 839 KLHLDLSLSPFASKPSPTAMILNLWEARHPNGNLSQLAAAVAGLQPDAGLFTVSEAC 898

## RESULT 5

US-09-970-944-13  
; Sequence 13, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herzman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shimkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1  
; TITLE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/237,862  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 898  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-09-970-944-13

Query Match 95.0%; Score 4545.5; DB 11; Length 898;  
Best Local Similarity 94.7%; Pred. No. 0;  
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;  
QY 1 MAVRPGIWPALLGIVLAALRGSGAQSATVANPVGANPDLLPHFLVEPEDVIVKNKP 60

```
Db 1 MAVRGLWFLVGLVLAAMLRGSGAQOSATVANVPFGANPDLLPHLVEPEDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120
QY 121 EYWCQCVAWSSGTTKSKQAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180
Db 121 EYWCQCVAWSSGTTKSKQAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180
QY 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
Db 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
QY 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
Db 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
QY 301 SWSWSKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDRNCTSDLCVHSASGPEV 360
Db 300 SWSWSKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDRNCTSDLCVHSASGPEV 359
QY 361 ALYVGLIAVAVCLVLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 420
Db 360 ALYVGLIAVAVCLVLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTITYQGSCLCPRODGPSPKFLTNGHLLSPILGGGRHTLHHSPTSEAEFV 479
Db 420 LTIQPDLS-TTITYQGSCLCRDGPSPKFLTNGHLLSPILGGGRHTLHHSPTSEAEFV 479
QY 480 SRLSTQNYFRSLPRGTSNNYTFNGLGRLMIPNTGISLILPPDAIPRGKIYEIYTLH 539
Db 480 SRLSTQNYFRSLPRGTSNNYTFNGLGRLMIPNTGISLILPPDAIPRGKIYEIYTLH 539
QY 540 KPEDVRLPLAGCQTLSPVSVCGPPGVLLTRPVILAMDHCGEPPDSWSLRLLKQCEGS 599
Db 540 KPEDVRLPLAGCQTLSPVSVCGPPGVLLTRPVILAMDHCGEPPDSWSLRLLKQCEGS 599
QY 600 WQDVULHGEABRSHIYYCOLEASACYVTEQGLRFPALVGEALSVAAARLKLILPAPVA 659
Db 600 WS-DVHLHGEESPSHLIYYCOLEAGACYVTEQGLRFPALVGEALSVAAARLKLILPAPVA 658
QY 660 CTSLEYNIRVYCHLHDTHDALKEVVQLEKQGLGQIQEPVLRHFKDSYHNLRSIHDPVSS 719
Db 659 CTSLEYNIRVYCHLHDTHDALKEVVQLEKQGLGQIQEPVLRHFKDSYHNLRSIHDPVSS 718
QY 720 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGDGQSF 779
Db 719 LWSKLLVSYQEIPIFYHIWNGTQRYLHCTFTLERSVSPSTDACKLWVWQVEGDGQSF 778
QY 780 NFNIKTDTFAELLALLESAGVPALVGPAPFKIPFLIROKIISLPPPCRRGADWRTLAQ 839
Db 779 NFNIKTDTFAELLALLESAGVPALVGPAPFKIPFLIROKIISLPPPCRRGADWRTLAQ 838
QY 840 KLHDSHLFFFAKSPPTAMILNLWEARHPNGLNSQLAAAVAGLGQPDAGLFTVSEAC 899
Db 839 KLHDSHLFFFAKSPPTAMILNLWEARHPNGLNSQLAAAVAGLGQPDAGLFTVSEAC 898
```

RESULT 6

US-10-256-702-5

Sequence 5, Application US/10256702

Publication No. US20030059859A1

GENERAL INFORMATION:

APPLICANT: Tessier-Lavigne, Marc

Leonardo, E. David

Hink, Lindsay

Masu, Masayuki

Kazuko, Keino-Masu

TITLE OF INVENTION: Netrin Receptors

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESS:

```
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 268 BUSH STREET, SUITE 3200
CITY: SAN FRANCISCO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/256,702
FILING DATE: 27-Sep-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/933,261
FILING DATE: 20-Aug-2001
APPLICATION NUMBER: 08/808,982
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: UC96-217
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 343-4341
TELEFAX: (415) 343-4342
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 898 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030059859A1 Relevant
TOPOLOGY: No. US20030059859A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-10-256-702-5
```

```
Query Match 95.0%; Score 4545.5; DB 14; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;
QY 1 MAVRGLWFLVGLVLAAMLRGSGAQOSATVANVPFGANPDLLPHLVEPEDEVYIVKNKP 60
Db 1 MAVRGLWFLVGLVLAAMLRGSGAQOSATVANVPFGANPDLLPHLVEPEDEVYIVKNKP 60
QY 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120
Db 61 VLLVCKAVPATQIFPKCNGEWRQVDHVIERTSDGSGSEPTMEVRINVSRQVQKVFGL 120
QY 121 EYWCQCVAWSSGTTKSKQAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180
Db 121 EYWCQCVAWSSGTTKSKQAYIRIARLKNFQEPLEKEVSLEQGVLCRPEGIPPAE 180
QY 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
Db 181 VEWLNRNEDLVDSLPNNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSAAVIVY 240
QY 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 300
Db 241 VNGGWSWTWESVCSASCGRWQKRSCTNPAPLNGGAFCEGQNV-QKTACATLCPVDG 299
QY 301 SWSWSKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDRNCTSDLCVHSASGPEV 360
Db 300 SWSWSKWSACGLDCTHWSRECSDPAPRNGGECOGTDLDRNCTSDLCVHSASGPEV 359
QY 361 ALYVGLIAVAVCLVLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 420
Db 360 ALYVGLIAVAVCLVLLLLALGLIYCRKKEGLSDVDSSILTSFGFQPSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTITYQGSCLCPRODGPSPKFLTNGHLLSPILGGGRHTLHHSPTSEAEFV 479
Db 420 LTIQPDLS-TTITYQGSCLCRDGPSPKFLTNGHLLSPILGGGRHTLHHSPTSEAEFV 479
```

```
QY 480 SRLSTONYFRSLPRGTSNMITYGTNFELGRLMIPNTGISLIIIPDPAIPRGKIYEIYTLH 539
Db 480 SRLSTONYFRSLPRGTSNMITYGTNFELGRLMIPNTGISLIIIPDPAIPRGKIYEIYTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQCEGS 599
Db 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQCEGS 599
QY 600 WEQDVHLHGEAPSHLYYQLEASACVYFTEQIGRFAVGEALSVAAAKRLKLLPAPVA 659
Db 600 WE-DVHLHGEAPSHLYYQLEAGACVYFTEQIGRFAVGEALSVAAATKRLRLPAPVA 658
QY 660 CTSLEYNIRVYCLHDTHDALKKEVQLEKQGGQIIOEPVLFHFKDSYHNLRLSIHDPVSS 719
Db 659 CTSLEYNIRVYCLHDTHDALKKEVQLEKQGGQIIOEPVLFHFKDSYHNLRLSIHDPVSS 718
QY 720 LWSKLLVSYQEIPIFYHWNQYRQYLHCTFTLERSVSPSTDACKLVWQVEGDGQSFSI 779
Db 719 LWSKLLVSYQEIPIFYHWNQYRQYLHCTFTLERSVSPSTDACKLVWQVEGDGQSFSI 778
QY 780 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIITASLDPPCSRGMWRTLAQ 839
Db 779 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIITASLDPPCSRGMWRTLAQ 838
QY 840 KLHLDSHLSFFASKPSPTAMILNLWEARHPFNGNLQAAAAVAGLQOPDAGLFTVSEAC 899
Db 839 KLHLDSHLSFFASKPSPTAMILNLWEARHPFNGNLQAAAAVAGLQOPDAGLFTVSEAC 898
```

## RESULT 7

US-10-240-154-16

; Sequence 16, Application US/10240154

; Publication No. US20030175741A1

; GENERAL INFORMATION:

; APPLICANT: Cochran et al.

; TITLE OF INVENTION: SCHIZOPHRENIA RELATED GENES

; FILE REFERENCE: CKFW-P01-006

; CURRENT APPLICATION NUMBER: US/10/240,154

; PRIOR FILING DATE: 2001-04-02

; PRIOR APPLICATION NUMBER: PCT/GB01/01486

; PRIOR FILING DATE: 2001-04-02

; NUMBER OF SEQ ID NOS: 34

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 16

; LENGTH: 898

; TYPE: PRT

; ORGANISM: Rattus sp.

US-10-240-154-16

```
Query Match 95.0%; Score 4545.5; DB 14; Length 898;
Best Local Similarity 94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3;

QY 1 MAVRPGMLPALLGIVLAAMLRGSGAQSATVANPFGANPDLLPHFLVPEPDIYVKNKP 60
Db 1 MAVRPGMLPALLGIVLAAMLRGSGAQSATVANPFGANPDLLPHFLVPEPDIYVKNKP 60
QY 61 VLLVCKAVPATQIFFKCNCEWVQDVHVIERTSDGSGSEPTMEVRINVSRRQVKEVFGLE 120
Db 61 VLLVCKAVPATQIFFKCNCEWVQDVHVIERTSDGSGSEPTMEVRINVSRRQVKEVFGLE 120
QY 121 EYWCQCVAWSSGTTKSKQAYIRIARLNKFNFEQPLAKEVSLRQGVLPKCRPEGIPPAE 180
Db 121 EYWCQCVAWSSGTTKSKQAYIRIARLNKFNFEQPLAKEVSLRQGVLPKCRPEGIPPAE 180
QY 181 VEWLRNEDLVDPGLDNNVYITREHSLVVRQARLADTANTYCAKNIVARRRSASAIVY 240
Db 181 VEWLRNEDLVDPGLDNNVYITREHSLVVRQARLADTANTYCAKNIVARRRSASAIVY 240
QY 241 VNGGWSTWTSVCSACSGRWKRSRSCNTNPAPLNGGAFCEGQNVHRTVSLVSDG 300
Db 241 VNGGWSTWTSVCSACSGRWKRSRSCNTNPAPLNGGAFCEGQNVHRTVSLVSDG 299
```

```
QY 301 SMSPSWKWACGLDCTHWRSECSDPAPRNGGEECCOGLDTRNCTSDLCVHSASGPEDV 360
Db 300 SMSPSWKWACGLDCTHWRSECSDPAPRNGGEECCOGLDTRNCTSDLCVHSASGPEDV 359
QY 361 ALYVGLIYAVAVCLVLLVLLVLYCYCKKEGLSDVADSSILTSFGQFVSIKPSKADNPHL 420
Db 360 ALYVGLIYAVAVCLVLLVLLVLYCYCKKEGLSDVADSSILTSFGQFVSIKPSKADNPHL 419
QY 421 LTIQPDLS-TTTTYQSLCPRQDPSPKFQLTNGHLLSPLGGRHTLHSSPSEAEFV 479
Db 420 LTIQPDLS-TTTTYQSLCPRQDPSPKFQLTNGHLLSPLGGRHTLHSSPSEAEFV 479
QY 480 SRLSTONYFRSLPRGTSNMITYGTNFELGRLMIPNTGISLIIIPDPAIPRGKIYEIYTLH 539
Db 480 SRLSTONYFRSLPRGTSNMITYGTNFELGRLMIPNTGISLIIIPDPAIPRGKIYEIYTLH 539
QY 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQCEGS 599
Db 540 KPEDVRLPLAGCOTLLSPVSCGPPGVLLTRPVILAMDHCGEPPDSWSLRLKKQCEGS 599
QY 600 WEQDVHLHGEAPSHLYYQLEASACVYFTEQIGRFAVGEALSVAAAKRLKLLPAPVA 659
Db 600 WE-DVHLHGEAPSHLYYQLEAGACVYFTEQIGRFAVGEALSVAAATKRLRLPAPVA 658
QY 660 CTSLEYNIRVYCLHDTHDALKKEVQLEKQGGQIIOEPVLFHFKDSYHNLRLSIHDPVSS 719
Db 659 CTSLEYNIRVYCLHDTHDALKKEVQLEKQGGQIIOEPVLFHFKDSYHNLRLSIHDPVSS 718
QY 720 LWSKLLVSYQEIPIFYHWNQYRQYLHCTFTLERSVSPSTDACKLVWQVEGDGQSFSI 779
Db 719 LWSKLLVSYQEIPIFYHWNQYRQYLHCTFTLERSVSPSTDACKLVWQVEGDGQSFSI 778
QY 780 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIITASLDPPCSRGMWRTLAQ 839
Db 779 NFNITKDTFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIITASLDPPCSRGMWRTLAQ 838
QY 840 KLHLDSHLSFFASKPSPTAMILNLWEARHPFNGNLQAAAAVAGLQOPDAGLFTVSEAC 899
Db 839 KLHLDSHLSFFASKPSPTAMILNLWEARHPFNGNLQAAAAVAGLQOPDAGLFTVSEAC 898
```

## RESULT 8

US-10-311-623-1

; Sequence 1, Application US/10311623

; Publication No. US20040023244A1

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.; GRIFFIN, Jennifer A.

; APPLICANT: KALLICK, Deborah A.; TRIBOULEY, Catherine M.

; APPLICANT: YUE, Henry; NGUYEN, Daniel B.

; APPLICANT: TANG, Y. Tom; LAL, Preeti G.

; APPLICANT: POLICKY, Jennifer L.; AZIMZAI, Valda

; APPLICANT: LU, Dyung Aina M.; GRAUL, Richard C.

; APPLICANT: YAO, Monique G.; BURFORD, Neil

; APPLICANT: HAFALIA, April J. A.; BAUGHN, Mariah R.

; APPLICANT: BANDMAN, Olga; ARVIZU, Chandra S.

; APPLICANT: YANG, Junming; XU, Yuming

; APPLICANT: GANDHI, Ameena R.; WARREN, Bridget A.

; APPLICANT: DING, Li; SANJANWALA, Madhusudan M.

; APPLICANT: DUGGAN, Brendan M.; LU, Yan

; TITLE OF INVENTION: RECEPTORS

; FILE REFERENCE: PF-0793 USN

; CURRENT APPLICATION NUMBER: US/10/311,623

; PRIOR FILING DATE: 2002-12-17

; PRIOR APPLICATION NUMBER: US 01/19942

; PRIOR FILING DATE: 2001-06-21

; PRIOR APPLICATION NUMBER: US 60/214,027

; PRIOR FILING DATE: 2000-06-21

; PRIOR APPLICATION NUMBER: US 60/228,045

; PRIOR FILING DATE: 2000-08-25

; PRIOR APPLICATION NUMBER: US 60/255,104

; PRIOR FILING DATE: 2000-12-12

; NUMBER OF SEQ ID NOS: 24

; SOFTWARE: PERL Program

```
; SEQ ID NO 1
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040023244A1 6052371CD1
US-10-311-623-1

Query Match      91.5%; Score 4379.5; DB 16; Length 842;
Best Local Similarity 93.1%; Pred. No. 0;
Matches 838; Conservative 1; Mismatches 2; Indels 59; Gaps 3;

QY 1 MAVRPLGIPALLIGVLAIAWLRGSGAQASATVANVPGANPDLLPHFLVPEDEVIVKNKP 60
DB 1 MAVRPLGIPALLIGVLAIAWLRGSGAQASATVANVPGANPDLLPHFLVPEDEVIVKNKP 60
QY 61 VLLVCCKAVPATQIFPKCKGEWVRQVDHVIERTDSSGSEPTWEVRINVSROQVEKVFGL 120
DB 61 VLLVCCKAVPATQIFPKCKGEWVRQVDHVIERTDSSGSEPTWEVRINVSROQVEKVFGL 120
QY 121 EYWCQVAMSSGGTTKSKAYIRIARLRKNFEQEPLAKEVSLQGIIVLPCRPEGIPPAE 180
DB 121 EYWCQVAMSSGGTTKSKAYIRIARLRKNFEQEPLAKEVSLQGIIVLPCRPEGIPPAE 180
QY 181 VEWLRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
DB 181 VEWLRNEDLVDPSPDNVYITREHSLVVRQARLADTANYTCVAKNIVARRRSASAIVY 240
QY 241 VNGGWSWTWMSVCSASCGRGWQKRSCTNPAPLNGGAFCEGQNVHDRTVSSLLVSDG 300
DB 241 -----VDG 243

QY 301 SWSPKWSACGLDCTHWSRECSDPAPRNGEECGTDLDRNCTSDLCVHSASGPEDV 360
DB 244 SWSPKWSACGLDCTHWSRECSDPAPRNGEECGTDLDRNCTSDLCVHTASGPEDV 303
QY 361 ALYVGLIAVAVCLVLLVLLVLYVCKKKEGLSDVADSSILTSQGPVSIKPSKADNPHL 420
DB 304 ALYVGLIAVAVCLVLLVLLVLYVCKKKEGLSDVADSSILTSQGPVSIKPSKADNPHL 363
QY 421 LTIQPDLS -TTTTYQGSCLPRQDGPSPKQLTNGHLLSPGGGRHTLHSSPTSABEFV 479
DB 364 LTIQPDLS -TTTTYQGSCLPRQDGPSPKQLTNGHLLSPGGGRHTLHSSPTSABEFV 423
QY 480 SRLSTQNYFRSLPRGTSNNMTYGTNFPLGGRLMIPNTGISLLIPDPAIPRGKIYEIYLT 539
DB 424 SRLSTQNYFRSLPRGTSNNMTYGTNFPLGGRLMIPNTGISLLIPDPAIPRGKIYEIYLT 483
QY 540 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 599
DB 484 KPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSCGS 543
QY 600 WEQDVLHGLEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAPVA 659
DB 544 WE -DVHLGLEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAPVA 602
QY 660 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDVPS 719
DB 603 CTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDVPS 662
QY 720 LMKSKLLVSYOEIPFYHIWNGTQRYLHCTFTLERVSPSTDACKLWVWQVEGDGQSF 779
DB 663 LMKSKLLVSYOEIPFYHIWNGTQRYLHCTFTLERVSPSTDACKLWVWQVEGDGQSF 722
QY 780 NFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDDPPCRRGADWRTLAQ 839
DB 723 NFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDDPPCRRGADWRTLAQ 782
QY 840 KLHLDLSHLSFFASKPSPMTAMINLWEARHPFNGNLSQLAATAVAGLQPDAGLFTVSEAC 899
DB 783 KLHLDLSHLSFFASKPSPMTAMINLWEARHPFNGNLSQLAATAVAGLQPDAGLFTVSEAC 842
```

```
RESULT 9
US-09-970-944-14
; Sequence 14, Application US/09970944
; Publication No. US20030204052A1
; GENERAL INFORMATION:
; APPLICANT: Heriman, John L
; APPLICANT: Rastelli, Luca
; APPLICANT: Shmukets, Richard A
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an
; FILE OF INVENTION: Antibodies Directed Against These Proteins
; FILE REFERENCE: 21402-138
; CURRENT APPLICATION NUMBER: US/09/970,944
; CURRENT FILING DATE: 2002-05-02
; PRIOR APPLICATION NUMBER: 60/237,862
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 544
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-970-944-14

Query Match      58.9%; Score 2819; DB 11; Length 544;
Best Local Similarity 99.6%; Pred. No. 4.2e-238;
Matches 540; Conservative 0; Mismatches 0; Indels 2; Gaps 2;

QY 359 DVALYVGLIAVAVCLVLLVLLVLYVCKKKEGLSDVADSSILTSQGPVSIKPSKADNP 418
DB 4 DVALYVGLIAVAVCLVLLVLLVLYVCKKKEGLSDVADSSILTSQGPVSIKPSKADNP 63
QY 419 HLLTIQPDLS -TTTTYQGSCLPRQDGPSPKQLTNGHLLSPGGGRHTLHSSPTSABEF 477
DB 64 HLLTIQPDLS -TTTTYQGSCLPRQDGPSPKQLTNGHLLSPGGGRHTLHSSPTSABEF 123
QY 478 FVSRSLSTQNYFRSLPRGTSNNMTYGTNFPLGGRLMIPNTGISLLIPDPAIPRGKIYEIYLT 537
DB 124 FVSRSLSTQNYFRSLPRGTSNNMTYGTNFPLGGRLMIPNTGISLLIPDPAIPRGKIYEIYLT 183
QY 538 LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSC 597
DB 184 LHKPEDVRLPLAGCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKKQSC 243
QY 598 GSWEQDVLHGLEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAP 657
DB 244 GSWE -DVHLGLEAPSHLYYQLEASACVYFTEQGRFALVGEALSVAALKLLFAP 302
QY 658 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDV 717
DB 303 VACTSLEYNIRVYCLHDTHDALKEVVQLEKQGGQLIQEPRVLHFKDSYHNLRLSIHDV 362
QY 718 SSLWSKLLVSYOEIPFYHIWNGTQRYLHCTFTLERVSPSTDACKLWVWQVEGDGQSF 777
DB 363 SSLWSKLLVSYOEIPFYHIWNGTQRYLHCTFTLERVSPSTDACKLWVWQVEGDGQSF 422
QY 778 SINFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDDPPCRRGADWRT 837
DB 423 SINFNITKDTFAELALAESEAGVPALVGPFAFKIPFLIRQKIISLDDPPCRRGADWRT 482
QY 838 AQKLLHLDLSHLSFFASKPSPMTAMINLWEARHPFNGNLSQLAATAVAGLQPDAGLFTVSEA 897
DB 483 AQKLLHLDLSHLSFFASKPSPMTAMINLWEARHPFNGNLSQLAATAVAGLQPDAGLFTVSEA 542
QY 898 EC 899
DB 543 EC 544

RESULT 10
US-09-933-261-6
; Sequence 6, Application US/09933261
; Publication No. US20030040046A1
```



QY 404 GFQVSIKPSKADNPHLLTIQDLS-TTTTVOGSLCPROQSPSPKFOITNGHLLSPGLGG 462  
Db 61 GFQVSIKPSKADNPHLLTIQDLS-TTTTVOGSLCPROQSPSPKFOITNGHLLSPGLGG 120  
QY 463 RHTLHSSPTSEAEFFVSRSTQNYFSLRGTSMNTYGTNEFLGGRLMIPNTGILLSIP 522  
Db 121 RHTLHSSPTSEAEFFVSRSTQNYFSLRGTSMNTYGTNEFLGGRLMIPNTGILLSIP 180  
QY 523 PDAIPRGKIYIYITLHKPEDVRLPAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEP 582  
Db 181 PDAIPRGKIYIYITLHKPEDVRLPAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEP 240  
QY 583 SPDGWSLRLKQSCGSEWQVHLGGEAPSHLYYQLEASACVYFTQOLGRFALVGEAL 642  
Db 241 SPDGWSLRLKQSCGSEW-DVLHLEGEAPSHLYYQLEASACVYFTQOLGRFALVGEAL 299  
QY 643 SVAARLKLILFAPVACTSLEYNIRVYCLDTHDALKKEVVQLEKQGGQILQIEPRVLHF 702  
Db 300 SVAARLKLILFAPVACTSLEYNIRVYCLDTHDALKKEVVQLEKQGGQILQIEPRVLHF 359  
QY 703 XDSYHNLRLSHDVPSSLWKSKLLVSYQIEIPFYHIWNGTORYLHCTFTLERSVSPSTDLA 762  
Db 360 XDSYHNLRLSHDVPSSLWKSKLLVSYQIEIPFYHIWNGTORYLHCTFTLERSVSPSTDLA 419  
QY 763 CKLWVQVEGQSQSFNINFTKTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIIS 822  
Db 420 CKLWVQVEGQSQSFNINFTKTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIIS 479  
QY 823 SLDPCCRGADWRTLAQKLHLDHLSFFASKPSPTAMIIINLWEARHFPNGLSQAAAVA 882  
Db 480 SLDPCCRGADWRTLAQKLHLDHLSFFASKPSPTAMIIINLWEARHFPNGLSQAAAVA 539  
QY 883 GLGQPDAGLFT-VSEAE 899  
Db 540 GTXPAGRWLLSQCSAE 557

RESULT 12  
US-09-970-944-15  
; Sequence 15, Application US/09970944  
; Publication No. US20030204052A1  
; GENERAL INFORMATION:  
; APPLICANT: Herrman, John L  
; APPLICANT: Rastelli, Luca  
; APPLICANT: Shinkets, Richard A  
; TITLE OF INVENTION: No. US20030204052A1el Proteins and Nucleic Acids Encoding Same an  
; TITLE OF INVENTION: Antibodies Directed Against these Proteins  
; FILE REFERENCE: 21402-138  
; CURRENT APPLICATION NUMBER: US/09/970,944  
; CURRENT FILING DATE: 2002-05-02  
; PRIOR APPLICATION NUMBER: 60/237,862  
; PRIOR FILING DATE: 2000-10-04  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 931  
; TYPE: PRT  
; ORGANISM: Caenorhabditis elegans  
US-09-970-944-15

Query Match 56.5%; Score 2705.5; DB 11; Length 931;  
Best Local Similarity 56.4%; Pred. No. 9e-228;  
Matches 515; Conservative 153; Mismatches 214; Indels 31; Gaps 12;

QY 9 PALLGIVLAWLKSGSAQQA---TVANPYGANPDLLPHLEVEPDVYIVKKNPVLLVC 65  
Db 26 PAL-ALLSASGTGSAQDDEFFHELPEFPDPPEPLHLEPEEAYIVKKNPNLYC 83  
QY 66 KAVPATQIFKCNSEWVRQVQVHVERSTGDSGSEPTMEVINSRQVQVKEVFGLEEYWCQ 125  
Db 84 KASPATQIFKCNSEWVRQVQVHVERSTGDSGSEPTMEVINSRQVQVKEVFGLEEYWCQ 143

QY 126 CVAWSSGTTKSKAYIRIARLNKFNQEPLEAKEVSLQGGIVLPCRPPEGIPPAEVEWLR 185  
Db 144 CVAWSSAGTTKSKAYIRIAYLRKTFEQPLKGEVSLQEVLLQCRPEGIPPAEVEWLR 203  
QY 186 NEDLVDPSPDPNVYITREHSLVVRQARLADTANTYCAKNIIVARRRSASAAVIVYVNGW 245  
Db 204 NEDIIDPAEDRNFYITIDHNLIIKQARLSDTANTYCAKNIIVARRRSASAAVIVYVNGW 263  
QY 246 STTWEMSVCSASCRGQKSRSCNTPAPLNGGAFCEGQNVHDTVSSLLVSDGSPW 305  
Db 264 STTWEMSVCSNCRGQKSRTRCTNTPAPLNGGAFCEGQSV-QKIACTTLCFVDRWTSW 322  
QY 306 SKWSACGLDCHMWSRECSDPAPRNGGEECCGTDLDTNCTSDLCVHSASGPEVDALVYG 365  
Db 323 SKWTCGTCTHWRRECTAPAPKNGGKCDGLVQSKNCTDGLCMQAPSDDDVALVYG 382  
QY 366 L-IAVAVCLVLLVLLVYCRKXEGDSDVADSSILTSQFQPSIYKPSKADNPHLLTIQ 424  
Db 383 IVIATVCLAITVVVALPVYRKNHRDPESDIIDSSALNGGFPVNIKAARQD---LLAVP 439  
QY 425 PDL-STTTTYQSLCPROQSPKFOITNGHLLSPGLGGRLTHHSS---PTSEAEFFV 479  
Db 440 PDLTSAAMYRGVVALHD-VSDKIPMTNPSILDLPLNKLKVYNSSGAVTPQDDLAEP 498  
QY 480 SRLS---TONYF-----RSLPRGT--SNMTYGTENFLGGRLMIPNTGILLSIP 525  
Db 499 SKLSPQMTQSLLENALNKLQSLARQDPSCTAPGTENSLGGHLIIIPNSGVLLIPAGA 558  
QY 526 IPRGKIYIYITLHKPEDVRLPAGCQTLSPVSCGPPGVLLTRPVILAMDHGCEPSPD 585  
Db 559 IPOGRVYEMVTVHRKENRPPMEDSQTLTPVWSCGPPGALLTRPVILTLLHHCADPSTE 618  
QY 586 SWSLELKKQSCGSEWQVHLGGEAPSHLYYQLEASACVYFTQOLGRFALVGEALSVA 645  
Db 619 DMKIQLKNQAVQGW-EVWVVGEEFTPCYIQLDAECHILTENLSTYALVGQSTTKA 677  
QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKKEVVQLEKQGGQILQIEPRVLHF 705  
Db 678 AAKRLKLAIFGLCCSSLEYSIRVYCLDDTDALKEVQLQERQMGQQLLEEFKALHFKGS 737  
QY 706 YHNLRLSHDVPSSLWKSKLLVSYQIEIPFYHIWNGTORYLHCTFTLERSVSPSTDACKL 765  
Db 738 IHNRLSLHDIHSLWKSKLLAKYQIEIPFYHWSGQRNLHCTFTLERSLNTVELVCKL 797  
QY 766 WTVQVEGQSQSFNINFTKTRFAELLALAESEAGVPALVGPSPAFKIPFLIRQKIIS 825  
Db 798 CVRQVEGQIQFLNCTVSEETPGIDPLDLPDASTITVTGSPSAFSPILPIRQKCSSLD 857  
QY 826 PPCRRGADWRTLAQKLHLDHLSFFASKPSPTAMIIINLWEARHFPNGLSQAAAVA 885  
Db 858 APQTRGHDRWRLAHKLNDRLYNYFATKSSPTGVILDLWEAQNFDPGNLSMLAAVLEEMG 917  
QY 886 QPDAGLFTVSEAB 898  
Db 918 RHETVTVSLAAEQ 930

RESULT 13  
US-10-087-684-35  
; Sequence 35, Application US/10087684  
; Publication No. US20040029116A1  
; GENERAL INFORMATION:  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Millet, Isabelle  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Stone, David J.  
; APPLICANT: Grosse, William M.  
; APPLICANT: Lepley, Denise M.  
; APPLICANT: Rieger, Daniel K.  
; APPLICANT: Burgess, Catherine E.  
; APPLICANT: Casman, Stacie, J.  
; APPLICANT: Spytek, Kimberly A.





QY	619	DMKIQIKNOAVQGWME-DVVVVGEENFTTPCYIQDABEACHILTENJSTYALVQGSTTKA	677
QY	646	AAKRLKLLIFAPVACTSLEYNIRVYCLDTHDALKVEVVOLEKQLGGOLIOBPVILHFKDS	705
Db	678	AAKRLKLAIFGLPCSSLEYSIRVCLDDTDOTALKEVLQLERQWGGQLLEFPKALHPKGS	737
QY	706	YHNRLSLTHDVPSSLWKSLLVSYQEIPIFYHWNQTRYLHCTFTLBERVSPSTDACKL	765
Db	738	IHNRLSLTHDIAHSLWKSLLAKYQEIPIFYHWSGQSNLHCTFTLBERLSLNTVELVCKL	797
QY	766	WYVQVGEQGSFNSINFNITKOTRPAELLALASEACVPALVGPSPAFKIPFLIROKLIISLSD	825
Db	798	CVRQVGEQGIQFLNCTVSEETGIDLLPLDPASTITTTVTGFSAFSIPLPRIQRKLCSSLD	857
QY	826	PPCRGADWRTLAOKLHLDLSLSPFASKPSPSTMTLNLEWARPNGNLISOLAAAVAGLG	885
Db	858	APQTRGHDMWLAHLNLDRLYNFYATKSSPTGVILDLWEAQNTPDGNLSMLAAVLEMG	917
QY	886	QPDAGLFTVSEAE	898
Db	918	RHETVTVSLAAEQG	930
RESULT 15			
US-10-037-417-117			
; Sequence 117, Application US/10037417			
; Publication No. US20040052806A1			
; GENERAL INFORMATION:			
; APPLICANT: Kekuda, Ramesh			
; APPLICANT: Alsobrook II, John P			
; APPLICANT: Tchernev, Velizar T			
; APPLICANT: Liu, Xiaohong			
; APPLICANT: Spytek, Kimberly A			
; APPLICANT: Patturajan, Meera			
; APPLICANT: Grosse, William M			
; APPLICANT: Lepley, Denise M			
; APPLICANT: Burgess, Catherine E			
; APPLICANT: Vernet, Corine A.M.			
; APPLICANT: Li, Li			
; APPLICANT: Gorman, Linda			
; APPLICANT: Edinger, Shlomit R			
; APPLICANT: Sciorre, Paul			
; APPLICANT: Ellerman, Karen			
; APPLICANT: Malvankar, Uriel M			
; APPLICANT: Rothenberg, Mark			
; APPLICANT: Stone, David J			
; APPLICANT: Boldog, Ferenc L			
; APPLICANT: Guo, Xiaojia			
; APPLICANT: Shenoy, Suresh G			
; APPLICANT: Anderson, David W			
; APPLICANT: Padigar, Muralidhara			
; APPLICANT: Taupier Jr, Raymond J			
; APPLICANT: Miller, Charles E			
; APPLICANT: Eisen, Andrew J			
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding: Same			
; FILE REFERENCE: 21402-235			
; CURRENT APPLICATION NUMBER: US/10/037,417			
; CURRENT FILING DATE: 2002-09-20			
; PRIOR APPLICATION NUMBER: 60/260,018			
; PRIOR FILING DATE: 2001-01-05			
; PRIOR APPLICATION NUMBER: 60/260,360			
; PRIOR FILING DATE: 2001-01-08			
; PRIOR APPLICATION NUMBER: 60/272,411			
; PRIOR FILING DATE: 2001-02-28			
; PRIOR APPLICATION NUMBER: 60/272,817			
; PRIOR FILING DATE: 2001-03-02			
; PRIOR APPLICATION NUMBER: 60/291,186			
; PRIOR FILING DATE: 2001-05-15			
; PRIOR APPLICATION NUMBER: 60/303,231			
; PRIOR FILING DATE: 2001-07-05			
; PRIOR APPLICATION NUMBER: 60/305,060			
; PRIOR FILING DATE: 2001-07-12			
; PRIOR APPLICATION NUMBER: 60/318,405			



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:23:31 ; Search time 50 seconds  
(without alignments)  
1729.524 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLMPALLGIVLAWL.....AVAGLQPDAGLFTVSEAC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	950.5	19.9	919	2 T32541	unc-5 protein - Ca
2	950.5	19.9	947	1 B44294	unc-5 protein, lon
3	280	5.8	1584	2 T00026	brain-specific ang
4	273.5	5.7	1172	1 TSHUP2	thrombospondin 2 p
5	273.5	5.7	1172	2 A42587	thrombospondin 2 p
6	270.5	5.7	1074	2 JC5928	semaphorin F precu
7	267.5	5.6	984	2 T00326	hypothetical prote
8	267.5	5.6	1522	2 T00028	brain-specific ang
9	263.5	5.5	1444	2 T18856	angiogenesis inhib
10	258	5.4	1572	2 T00027	brain-specific ang
11	248	5.2	1170	2 A40538	thrombospondin 1 p
12	246	5.1	1170	1 TSHUP1	thrombospondin 1 p
13	243.5	5.1	1178	1 A39804	thrombospondin pre
14	220.5	4.6	1651	2 T14160	transmembrane rece
15	217.5	4.5	1612	2 T30805	cutti protein - mo
16	211.5	4.4	1265	1 A37967	neural cell adhesi
17	208.5	4.4	469	1 S29126	properdin precurs
18	206.5	4.3	254	2 T15952	hypothetical prote
19	202.5	4.2	437	2 S05478	properdin - mouse
20	198.5	4.1	788	2 T25061	hypothetical prote
21	188	3.9	423	2 T29549	hypothetical prote
22	185	3.9	1344	2 T14316	hypothetical prote
23	184	3.8	957	2 T15976	rig-1 protein - mo
24	183.5	3.8	1273	2 T42405	hypothetical prote
25	183.5	3.8	1863	2 S46217	sax-3 protein - Ca
26	176.5	3.7	860	2 T16892	protein-tyrosine-p
27	175.5	3.7	1736	2 A47747	hypothetical prote
28	170.5	3.6	1745	2 A46431	tight junction pro
29	169	3.5	934	1 A34372	complement C6 prec

30	167.5	3.5	1907	2 S50893	protein-tyrosine-p
31	166	3.5	837	2 T00355	hypothetical prote
32	160.5	3.4	550	2 T47158	hypothetical prote
33	152.5	3.2	1499	2 I50212	protein-tyrosine-p
34	152	3.2	951	2 T00017	gene ADAMTS-1 prot
35	152	3.2	1501	2 I58148	protein-tyrosine-p
36	151	3.2	2165	2 T21371	hypothetical prote
37	150	3.1	1277	2 T30532	neural cell adhesi
38	149	3.1	654	2 T29247	hypothetical prote
39	148.5	3.1	946	1 A47299	ror-related recept
40	148	3.1	1091	2 A58532	glial cell membran
41	147	3.1	843	1 A27340	complement C7 prec
42	147	3.1	888	2 S23065	ufo protein - mous
43	146.5	3.1	584	1 C8HUA	kinase-like protei
44	146	3.0	1051	2 A39712	protein-tyrosine-p
45	145.5	3.0	2029	1 TDFFLK	

## ALIGNMENTS

### RESULT 1

T32541

unc-5 protein - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 31-Jan-2000

C:Accession: T32541

C:Ratrefille, P.

submitted to the EMBL Data Library, December 1997

A:Description: The sequence of C. elegans cosmid B0273.

A:Reference number: Z21187

A:Accession: T32541

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-919 <LAT>

A:Cross-references: EMBL:AF036698; PIDN:AB88355.1; GSPDB:GN00022; CESP:B0273.4a

A:Experimental source: strain Bristol N2; clone B0273

C:Genetics:

A:Gene: unc-5; CESP:B0273.4a

A:Map position: 4

A:Introns: 41/3; 108/1; 142/3; 201/1; 323/2; 553/1; 859/3

C:Superfamily: unc-5 protein; immunoglobulin homology; SH3 homology; thrombospondin type

Query Match 19.9%; Score 950.5; DB 2; Length 919;  
Best Local Similarity 28.5%; Pred. No. 1.6e-61;  
Matches 265; Conservative 165; Mismatches 375; Indels 125; Gaps 31;

QY 49 EPEDVYTKNPKVLLVCKAVPATQIFPKNGEWWVQVDHVIER--STDGSSGEPTMEVRI 106  
Db 9 QPKSGYVIRNKPRLQCRANHATKIRYKCSKWID--DSRIEKLIGTDSTSGVGVIDASV 66

QY 107 NVSRQQVEKVFGLBEYWCQCQVAVSSSG-----TTKSQKAVIRIARLRNFECEPLAKEVS 161  
Db 67 DISRIDVDTSGHVDAFQCQCYA---SGDDQDVVASDVATVHLAYMKHFLKSPVAQRVQ 123

QY 162 LEQGIIVLPCPPGPIPPAEVWELRNEDLVPDLDPNVIITREHSLVVVRQRLADTANYTC 221  
Db 124 EGTTLQPCAPESDPKAEITWKDGVVQVP--DANVIRASDGLISWAARLSDSGNVTC 181

QY 222 VAKNIVARRRSASAAVIVVGGWSTWTEW-SVCSASCG-----RGQKR 265  
Db 182 EATNVANSRATDPVEQIYVDGGSWSPWIGTCHVDCPLLRQHAHRIRDPHDLPHQRR 241

QY 266 SRSTCTNPALNGAFCEGQNVHRTVSSLLVSDGWSWMSKWSACGLDCTHWSRECS 325  
Db 242 TRICNNPAPLNDGEYCKGEBEMTRS-CKVCKLDGGSSWSWSSACSSCHRRVTRACTV 300

QY 326 PAPNGEECCQGTDLTRNCTSLCVHVSAG--PEDVALYVGLIAVAVCLVLLVLLV 383  
Db 301 PPMNGGQPCFGDMLTQCEPAQLCTADSRIVISDTAVYGSVASIFIVASFILAILAMF 360

QY 384 YCR-----KKEGLSDVADS-----SILTSGFQ 407

361 CCKRGSKSKPLPKQKMNSEKAGGIYYSEPPGVRELLLEHQHGHTLLGEKISSCSQYFEP 420  
408 VSIKPSKADNPHLLTIQPDLSITTTYYQGSICRQDQSPSKFQLTNGHLLSPGCGRHLLH 467  
421 PPL-----PHSTTLR---SGKSAPSGYSYTRNAGSRAALIOECSSSSSGSGKRTNLR 470  
468 HSSPT-SEAEFVSRLSTQNYFRSLPRGYS-NMTYGTFTNPLGRLMIPNTGTISILLPPDA 525  
471 TSSNCSDDNDYATLDYMEDKSVLGLDTSONIVAAQIDNSGARLSLSKSGARLIVPELA 530  
526 IPRGKIYIYTLHKPEDVRLPLAGCOTLLSPIVSCGPGCV-----LLTRPVILAMDH 578  
531 VEGEKM--LXLAVSDTLTDQPHLKPTEALSPIVIVICQCDVMSAHDNIILRRPVVVSFRH 588  
579 CGEESP-DSWSLRLLKKQSCGEGSWEODVLHGREAPSHLYYCOLEASA-----CVVFT 629  
589 CASTFPRDNWQFTL--YADEGSGWQKAVTIGEENLNTNMFVQEPQPKKNDGFGWCHMT 646  
630 EQLGRFALVGEAL--SVAAAKLLKLLFAPVACTSLE--YNIIRVYCLHTDHALKEVVQL 685  
647 YSLARLMLAGHPRNLSLSAAKRVHLAVFGPTENSAYRRPELRYVCVPETGAAMESVMKQ 706  
686 EKQLGGLTIQEP--VHLPGKDSYHNRLSLIHDV-PSLWKSLLVSVQRIFFYHNGTQ 742  
707 ED--GSRLLCESNDFILNEKG--NLCICIEDVIPGSCDGPEVVEISIQHRFV---AQ 758  
743 RYLHCTFLIRVSPSTSDIACKLWVQVGDGQSFNFINITKDTFAELLALAESEAGVP 802  
759 NGLHCSLFRPKBEINGSQFSTRVIVYQKASSTFPMVN--EVSNEPELYDATSEREKGSV 816  
803 ALVGPSAFKIPFLIRKIISSIDPPCRRRGADWRTLAQKLHLDLSLFFASKP--SPTAMI 860  
817 CV----EFLRPGVKDELARLLDMPNESHSDWRGLAKKLHYDRYLQFFASFDCSPSLL 872  
861 LNLWFAHPFNGN-LSQLAAAVAGLGOPDA 889  
873 LDLWEASSSGSARAVPDLQLTLRVMGPRDA 902

RESULT 2  
B44294  
unc-5 protein, long form - *Caenorhabditis elegans*  
N:Contains: unc-5 protein, short form  
C:Species: *Caenorhabditis elegans*  
C>Date: 30-Apr-1993 #sequence\_revision 28-Jul-1995 #text\_change 05-Nov-1999  
C:Accession: B44294; T32540; A44294  
R:Leung-Hageteitj, C.; Spence, A.M.; Stern, B.D.; Zhou, Y.; Su, M.W.; Hedgecock, E.M.;  
Cell 71, 289-299, 1992  
A:Title: UNC-5, a transmembrane protein with immunoglobulin and thrombospondin type 1 do-  
A:Reference number: A44294; MUID:93046629; PMID:1384987  
A:Contents: variety Bergerac  
A:Accession: B44294  
A:Molecule type: DNA  
A:Residues: 1-947 <LEU>  
A:Cross-references: GB:S47168; NID:G258527; PIDN:AA23867.1; PID:G258529  
A>Note: sequence extracted from NCBI backbone (NCBIN:116668, NCBIN:116670, NCBIN:116672,  
A>Note: authors translated the codon CTA for residue 642 as Val; sequence shown follows  
A>Note: mRNA lacking the first exon is equally prevalent  
R:Latreille, P.  
submitted to the EMBL Data Library, December 1997  
A:Description: The sequence of *C. elegans* cosmid B0273.  
A:Reference number: Z1187  
A:Accession: T32540  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-947 <LAT>  
A:Cross-references: EMBL:AF036658; PIDN:AA88356.1; GSPDB:GN00022; CESP:B0273.4b  
A:Experimental source: strain Bristol N2; clone B0273  
C:Genetics:  
A:Gene: unc-5  
A:Map position: 4  
A:Introns: 28/1; 69/3; 136/1; 170/3; 229/1; 351/2; 581/1; 886/3  
A:Function:

Qy	803	ALVGPSAEKIP	ELIRBQKLI	SSLDPPCR	CGADNR	TTIAQK	LHLSHLS	FFASKP	---SPTAM	860
Db	845	CV----	EPRLPGV	KDELAR	LMDNESH	DSMDRGL	AKKLHYR	LYQFFAS	FPDCPS	900
Qy	861	LNLWEAR	HFPGN	-LSQ	AAAVAG	QGPDA	889			
Db	901	LDLWEA	SSGSGA	RAVP	DLQLT	LRV	VMGR	PD	930	

### RESULT 3

T00026  
brain-specific angiogenesis inhibitor 1 - human  
N/Alternate names: BAIL protein  
C/Species: Homo sapiens (man)  
C/Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 12-Feb-1999  
C/Accession: T00026  
R/Nishimori, H.; Shiratsuchi, T.; Urano, T.; Kimura, Y.; Kiyono, K.; Tatsumi, K.; Yoshida, K. et al. (1999) The BAIL protein, a novel member of the thrombospondin type 1 repeat homology family, is a secreted protein that induces angiogenesis in vitro and in vivo. *J Biol Chem* 274:11507-11514  
submitted to the EMBL Data Library, June 1997  
A/Reference number: Z14064  
A/Accession: T00026  
A/Status: translated from GB/EMBL/DDBJ  
A/Molecule type: mRNA  
A/Residues: 1-1584 <NIS>  
A/Cross-references: EMBL:AB005297; NID:d1175078; PID:d1024528  
A/Experimental source: brain  
C/Genetics:  
A/Gene: GDB:BAIL  
A/Cross-references: GDB:9838088; OMIM:602682  
A/Map position: 8q24-8q24  
C/Superfamily: thrombospondin type 1 repeat homology  
C/408-462/Domain: thrombospondin type 1 repeat homology <THR3>

	Query Match	5.8%;	Score 280;	DB 2;	Length 1584;
	Best Local Similarity	31.8%;	Pred. No. 6.5e-12;		
	Matches	74;	Conservative	38;	Mismatches 93;
					Indels 28;
					Gaps 10;
QY	124	CQCWAMSSGTTKSQAKYIARIARKFPEOPLAKEVSLGQIVLPCRPPEGIPPAEVEW	183		
Db	309	CNREACGPAGTSSRSOSLSSTDARR---REELGDEL---QQGFPPA-PQTGDDPAAE-EW	360		
QY	184	LRNEDLVDPDLPNVYITREHSLVVRQARLADTANYTCVAKNIIVARERSASAAVIVVNG	243		
Db	361	--SPWVSCSTCGEGWQTR-----TRFCVSSSYSTQCSGPLREQLCNNSAVCPVHG	410		
QY	244	GWSTWTWVSCSASCGRGWOKRSCTNPALNGACFEGONVHDRVSSLV---SVDG	300		
Db	411	AWDEWFWSLCSSTCGGFRDRTNTR--PPQFGNPECEKQTKFCNTALCEGRAVDG	468		

## RESULT 4

TSHUP2  
Thrombospondin 2 precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 19-May-1995 #sequence\_revision 03-Aug-1995 #text\_change 13-Aug-1999  
C:Accession: A47379; A42173  
R:LaBell, T.L.; Byers, P.H.  
Genomics 17, 225-229, 1993  
A:Title: Sequence and characterization of the complete human thrombospondin 2 cDNA: pote  
A:Reference number: A47379; MUID:94010892; PMID:8406456  
A:Accession: A47379  
A:Molecule type: mRNA  
A:Residues: 1-1172 <LAB>  
A:Cross-references: GB:L12350; NID:g307505; PIDN:AAA03703.1; PID:g307506  
R:LaBell, T.L.; Milewicz, D.J.; Distech, C.M.; Byers, P.H.  
Genomics 12, 421-429, 1992  
A:Title: Thrombospondin II: partial cDNA sequence, chromosome location, and expression c  
A:Reference number: A42173; MUID:92217961; PMID:1559694  
A:Accession: A42173  
A:Molecule type: mRNA

A;Residues: 560-1172 <LA2>  
A;Cross-references: GB:M81339  
A;Experimental source: fibroblast  
A;Note: sequence extracted from NCBI backbone (NCBIN:95091, NCBI:P:95096)  
C;Genetics:  
A;Gene: GDB:THB52; TSP2  
A;Cross-references: GDB:128789; OMIM:188061  
A;Map position: 6q27-6q27  
C;Complex: homotrimer, disulfide linked  
C;Function:

A:Description: participates in cell migration and adhesion, and in platelet aggregation  
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; von  
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trimer  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-1172/Product: thrombospondin 2 #status predicted <MAT>  
F:319-377/Domain: von Willebrand factor type C repeat homology <VWC>  
F:380-431/Domain: thrombospondin type 1 repeat homology <THR1>  
F:436-492/Domain: thrombospondin type 1 repeat homology <THR2>  
F:493-549/Domain: thrombospondin type 1 repeat homology <THR3>  
F:553-588/Domain: EGF homology <EGF1>  
F:652-691/Domain: EGF homology <EGF>  
F:928-930/Region: cell attachment (R-G-D) motif  
F:151,316,330,457,584,710,1069/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:167-226/Disulfide bonds: #status predicted  
F:266,270/Disulfide bonds: interchain #status predicted  
F:612/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted

Query Match	5.7%;	Score 273.5;	DB 1;	Length 1172;
Best Local Similarity	28.9%;	Pred. No. 1.3e-11;		

209	QY	QBARLADTANYTCVAKNIIVARRASAA-VIVYVNGGSGTWTSVCSASCGRGWQKESR	267
403	Db	QGRSCDVTSNTCLGPSIQTRACSLSKCDTRIQDGGHSHSPWSSCSVTCGVGNITRIR	462
268	QY	SCTNPAFLNGAFCGQGVHDRTVSSLLVSDGSGSPWSKWSACGLDCT---HWESRCS	324
463	Db	LCNSFPQMGKGNCKSGRETKACQAGPCPDIGRWSPWSACTVTCAGGIRBTRVCN	522
325	QY	DPAPRNGEGCGTDLDRNCTSDLCVHSASGPEDEVALYVGLIAVAVCLVLLLVILVY	384
523	Db	SPEFQYGGKACGVQVQBRQMCNRSC-----PVDGCLSNPCFFGAQC-	564
385	QY	CRKKEGLSDVADSSILTSGFQPSVI--KPSKADNPHLITQPLDLSITT-----TYQ	434
565	Db	-----SSFPDGS-WSCGFCPVGLNGHCHCEDLDECALVPDICTFSTSKVPCVNTQP	615
435	QY	GSIC-----PRQDGPSP	446
616	Db	GFHCLPCPPRYRGNQP	631

## RESULT 5

A42587  
Thrombospondin 2 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change  
C:Accession: A42587; A39851  
R:Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.  
J. Biol. Chem. 267, 3274-3281, 1992  
A:Title: Characterization of mouse thrombospondin 2 sequence an  
A:Reference number: A42587; MUID:92147683; PMID:1371115  
A:Accession: A42587  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-1172 <LAH>  
A:Cross-references: GB:L07803; GB:M87275; NID:G340421; PIDN:AAA  
A:Note: sequence extracted from NCBI backbone (NCBI:P:81502)  
R:Bornstein, P.; O'Rourke, K.; Wikstrom, K.; Wolf, F.W.; Katz, R.  
J. Biol. Chem. 266, 12821-12824, 1991  
A:Title: A second, expressed thrombospondin gene (Thbs2) exists  
A:Reference number: A39851; MUID:91302287; PMID:1712771  
A:Accession: A39851



## RESULT 9

T18856  
angiogenesis inhibitor homolog - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
C:Accession: T18856, T24653  
R:McMurray, A.  
submitted to the EMBL Data Library, July 1995  
A:Reference number: Z19031  
A:Accession: T18856  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1444 <W1>  
A:Cross-references: EMBL:Z50004; PIDN:CAA90293.1; GSPDB:GN00028; CESP:C02B4.1  
A:Experimental source: clone C02B4  
R:McMurray, A.  
submitted to the EMBL Data Library, July 1995  
A:Reference number: Z19917  
A:Accession: T24653  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1444 <W12>  
A:Cross-references: EMBL:Z50006; PIDN:CAA90302.1; GSPDB:GN00028; CESP:C02B4.1  
A:Experimental source: clone T07C5  
C:Genetics:  
A:Gene: CESP:C02B4.1  
A:Map position: X  
A:Introns: 25/3; 70/3; 96/3; 139/3; 187/1; 234/2; 282/3; 376/2; 422/2; 478/3; 509/3; 568

Query Match 5.5%; Score 263.5; DB 2; Length 1444;  
Best Local Similarity 26.4%; Pred. No. 9.3e-11;  
Matches 73; Conservative 26; Mismatches 91; Indels 87; Gaps 12;  
QY 123 WCOCVWSSSGTTKSKAYIRIARLRKNFEQ-----EPLAKVSVLEQGVLPFCRPEGI 176  
DB 1134 WSEWSSNSAC-----SCFSLSTRRRFCQVVDPTVQGPCAGAILQ-----IFCAPGSCS 1183  
QY 177 PPAAE-----VEW-----LRNEDLVPSLDPNVYITREHSLVVRQARLADTAN 218  
DB 1184 PSAGGWSLSEWSSCKDCGDTGHQIRNRCSEF-----IPSNRGAYCSG 1228  
QY 219 YT-----CYAKNIVARRRSASAAVYVNGSGWSTWESVCSASCGRGWOKRSRCTNPA 273  
DB 1229 YSFQRPVMDNVCSDEK-----VDGGWTDWTAWSECTDYCRNGHRSRTRFCANPK 1279  
QY 274 PLMGGAFCGQN-----VHRTVSSLLVSDGWSWPSKWSACGLDC---THWR 319  
DB 1280 PSQGGHACTGSDPELNPFCPPARCHLR-----DGGWSTWSDWTPCSASC GFVQTR 1330  
QY 320 SRECSDPAPNGGEECGTDLDRNCTSDLCVHSASG 356  
DB 1331 DRSCSSPEPK-GGQSCSGLAHQTSCLDLPACDHS DG 1366

## RESULT 10

T00027  
brain-specific angiogenesis inhibitor 2 - human  
N:Alternate names: BAI2 protein  
C:Species: Homo sapiens (man)  
C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000  
C:Accession: T00027  
R:Shiratsuchi, T.; Nishimori, H.; Ichise, H.; Nakamura, Y.; Tokino, T.  
Cytogenet. Cell Genet. 79, 103-108, 1997  
A:Title: Cloning and characterization of BAI2 and BAI3, novel genes homologous to brain  
A:Reference number: Z14066; MUID:98194217; PMID:9533023  
A:Accession: T00027  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1572 <SHI>  
A:Cross-references: EMBL:AB005298; NID:g3021698; PIDN:BAA25362.1; PID:g3021699  
A:Experimental source: brain  
C:Genetics:

A:Gene: GDB:BAI2  
A:Cross-references: GDB:9838089; OMIM:602683  
A:Map position: 1p35-1p35

Query Match 5.4%; Score 258; DB 2; Length 1572;  
Best Local Similarity 18.8%; Pred. No. 2.6e-10;  
Matches 172; Conservative 110; Mismatches 312; Indels 322; Gaps 37;

QY 173 PEGIPPAEVEWURNEDLVDPSPDNVY-----ITREHSLVVRQARL 213  
DB 271 PEEEPKVKTPWPSAD-----EPGLYMAQTGDPAAEWSWPSVCSLTGCGQLQVR-TKS 323  
QY 214 ADTANYTCVAKNIVARRRSASAAVYVNGSGWSTWESVCSASCGRGWOKRSRCTNPA 273  
DB 324 CVSSPYGLCSPLRETRPCNNSATCPVHGWEWGSNSLCSRSCGRGSRMRCTCV--P 381  
QY 274 PLNGGAFCEGQNVHRTVSSLLVSDGWSWPSKWSACGLDC---THWRSRECSDPAPR- 329  
DB 382 PQHGGKACEGPQLQTKLCSMAACPVEGQWLEWGPWGPCSTSCANGTQQRSKCSVAGPAW 441  
QY 330 -----NGBEECQ 336  
DB 442 ATCTGALTDTRECSNLCEPATDSKMGPMNWSLCSKTCDTGMQRFRMCOATGTQYPCE 501  
QY 337 GTDLDRNCTSDLC--VHSASGPELVAL----- 362  
DB 502 GTGEEVKPCSEKRCFAPHEMCRDEYVMLMTWKAAGEIYINKPCPNASGSASRRCLLSA 561  
QY 363 ----YVGLHAAVCL---VLLLVLLVYCRKKEGSDVDVADSSILTSGFQPVSIKPSKA 415  
DB 562 QGVAYWGLPSFARCI SHEYRYLYLSREHLAKQRMAGEKGSQVVRSLQELLARRTYY 620  
QY 416 DNPILLTIQPDLSITTTY-QGSLCPRQDPSPKFOLT-----NGHLLPLGG 461  
DB 621 SGDLLEFSDILRNVDITFKRATVYPSADVDVRFQVVSFPMVDANKKMDDAQVSP--G 678  
QY 462 GRHTLHSSPTSEAEFV-----SRLSNTQYFRSLPRG-----TSNMTYGTEN 504  
DB 679 SVHLR-----VVEFIIHLVGDALAKAFQSSLLVTNLSIQREPVSAVSSDITPMRG 732  
QY 505 FLG-----GRLMIPNTGISLLIP-----PDALPRGK----- 530  
DB 733 RRGMKDWVRHSEDRFLPKPEVLSLSPKPAATSGAAGSPGRGPGTVPDPGSHQRL 792  
QY 531 -----IYE-IYLTILHKPEDVRULPACQTLISPIVSCGPGVLLTRPVIL 574  
DB 793 PADPDESSYFVIGAVLYRTILGLILPPP---RPPLAVTSRYMT--VTVRPPTQPPAEPLIT 847  
QY 575 A-----MDHCGEPPDSWSLRKQSCGSEWQDVLHGEAPSHLYYCOLEASACVFT 629  
DB 848 VELSYIINGTTDPHCASWDYS-RADASSGDWDNTENCOTLETOAHT-RCQ-----C 896  
QY 630 EQLGRFALYGE-----ALSVAARLKLILLFAPVACTSLEYNIRVYCLHDTDALKEVVQ 684  
DB 897 QHLSLFAVLAQPPKDLTLELAGSPSVPLVIGCAVSCWALLTLAIYA-----AFWRFIK 950  
QY 685 LEKQLGQLIQEPRVLHFKDSYHNRLSIHDVPSSLWKSLLYSYQEIYPYHLWNGTQRY 744  
DB 951 SERSI-----ILLNFCLSI--LASNI---LILVQSRVLSKGVCTMTAAF 990  
QY 745 LHCTFTLERVSPSTSLACKLMV-----WQVEGDGQS 776  
DB 991 LHFFF-----LSSFCWLTWQSYLAVIGEMRTRLYRKRFLCLGWLGPALVVA 1039  
QY 777 FSINFNITKDTRFAPALLALESEAG-VPALVGPQA-----FKTFFLIROKI-----ISSL 824  
DB 1040 VSVGFTRTKGYGFSYCWLSLEGGLLYAFVGPAAVILVNMNLGIIIVFNKLMARDGISDK 1099  
QY 825 DPPCRGAD---WRTL 837  
DB 1100 SKQKAGASERCPWASL 1115



RESULT 11

A40558  
thrombospondin 1 precursor - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 05-Jun-1992 #sequence revision 05-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40558; A37905; B42587; S68787  
R:Lawler, J.; Duquette, M.; Ferro, P.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.  
Genomics 11, 587-600, 1991  
A:Title: Characterization of the murine thrombospondin gene.  
A:Reference number: A40558; MUID:921128941; PMID:1774063  
A:Accession: A40558  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1170 <LAW>  
A:Cross-references: GB:M62449; GB:M62450; GB:M62451; GB:M62452; GB:M62453; GB:M62454; GB:M62465; GB:M62466; GB:M62467; GB:M62468; GB:M62469; GB:M62470; NID:G511867; PIDN:AAA50616  
R:Bornstein, P.; Alfai, D.; Devarayalu, S.; Franson, P.; Li, P.  
J. Biol. Chem. 265, 16691-16698, 1990  
A:Title: Characterization of the mouse thrombospondin gene and evaluation of the role of  
A:Reference number: A37905; MUID:90375546; PMID:2398070  
A:Accession: A37905  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-490 <BOR>  
A:Cross-references: GB:J05605; GB:J05606; NID:G201991; PIDN:AAA40431.1; PID:G554390  
R:Laherty, C.D.; O'Rourke, K.; Wolf, F.W.; Katz, R.; Seldin, M.F.; Dixit, V.M.  
J. Biol. Chem. 267, 3274-3281, 1992  
A:Title: Characterization of mouse thrombospondin 2 sequence and expression during cell  
A:Reference number: A42587; MUID:92147683; PMID:1371115  
A:Accession: B42587  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-1152 'p', 1154-1170 <LAH>  
A:Cross-references: GB:M87276  
A:Note: sequence extracted from NCBI backbone (NCBIP:81501)  
R:Chen, H.; Aeschlimann, D.; Nowlen, J.; Mosher, D.F.  
FEBS Lett. 387, 36-41, 1996  
A:Title: Expression and initial characterization of recombinant mouse thrombospondin 1 a  
A:Reference number: S68787; MUID:96234006; PMID:8654563  
A:Accession: S68787  
A:Molecule type: protein  
A:Residues: 19-26, 'X', 28-37 <CHE>  
C:Complex: homotrimer, disulfide linked  
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v  
C:Keywords: calcium binding; glycoprotein; homotrimer  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-1170/Product: thrombospondin 1 #status predicted <MAT>  
F:317-375/Domain: von Willebrand factor type C repeat homology <VNC>  
F:378-429/Domain: thrombospondin type 1 repeat homology <THR1>  
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>  
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>  
F:551-586/Domain: EGF homology <EGF>  
F:248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	5.2%;	Score 248;	DB 2;	Length 1170;
Best Local Similarity	29.9%;	Pred. No. 9.5e-10;		
Matches	53;	Conservative 24;	Mismatches 76;	Indels 24; Gaps 4;

  

QY	207	VYRQRLADTANYTCVAKNIIVAR-----RRSASAAVIVYVNGSWGTWTSWCSASC	258
dDb	399	IQQGRSCDSLNNRCGEGSVQTRTCHI QECDXRFKQ-----DGGWHSWSPWSCSVTC	451
QY	259	GRGWOKRRSCTNPAPLNGGAFCEGQNVHDIRTVSSLLVSDGWSWSPWKSWSACGLDC---	315
dDb	452	GDGVITRIRLCNSPSPQWNGKPCGEARETRACKDKDACPINGCWGFWPMDICSVTCGGG	511
QY	316	THWSRSCSDPAPRNGGEECOGTDLDTNRCTSDLCVHSASGPEDEVALYVGLIAVAVC	372
dDb	512	VQRRSLRCLNPTPOPGGKDCVGDVTENQVCMKQDC-----PIDCLSNPCFAGAKC	562

RESULT 12  
TSHUP1

thrombospondin 1 precursor - human

C:Species: Homo sapiens (man)  
C>Date: 23-Aug-1987 #sequence, revision 03-Aug-1995 #text change 17-Nov-2000  
C:Accession: A26155; A34274; A30140; A25812; A05172; A42927  
R:Lawler, J.; Hynes, R.O.  
J. Cell Biol. 103, 1635-1648, 1986

A:Title: The structure of human thrombospondin, an adhesive glycoprotein with multiple  
A:Reference number: A26155; MUID:87057617; PMID:2430973  
A:Accession: A26155  
A:Molecule type: mRNA  
A:Residues: 1-1170 <LA>  
A:Cross-references: GB:J04835; PIDN:CAA28370.1; PID:g37138  
A>Note: parts of this sequence, including the amino end of the mature protein, were det  
R:Jaherty, C.D.; Gierman, F.M.; Dixit, V.M.  
J. Biol. Chem. 264, 11222-11227, 1989

A:Title: Characterization of the promoter region of the human thrombospondin gene. DNA  
A:Reference number: A34274; MUID:89291870; PMID:2544587  
A:Accession: A34274  
A:Molecule type: DNA  
A:Residues: 1-166 <LAH>  
A:Cross-references: GB:J04835  
R:Hennessy, S.W.; Frazier, B.A.; Kim, D.D.; Deckwerth, T.L.; Baumgartel, D.M.; Rotwein,  
J. Cell Biol. 108, 729-736, 1989

A:Title: Complete thrombospondin mRNA sequence includes potential regulatory sites in th  
R:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.  
Biochemistry 25, 8418-8425, 1986

A:Title: Partial amino acid sequence of human thrombospondin as determined by analysis o  
A:Reference number: A25812; MUID:87157592; PMID:3030396  
A:Accession: A25812  
A:Molecule type: mRNA  
A:Residues: 1-83, 'A', 85-522, 'A', 524-1170 <HEN>  
A:Cross-references: EMBL:X14787; NID:G37464; PIDN:CAA32889.1; PID:G37465  
A>Note: parts of this sequence, including the amino end of the mature protein, were det  
R:Kobayashi, S.; Eden-McCutchan, F.; Framson, P.; Bornstein, P.  
Biochemistry 25, 8418-8425, 1986

A:Title: Partial amino acid sequence of human thrombospondin as determined by analysis o  
A:Reference number: A05172; MUID:86287276; PMID:3461443  
A:Accession: A05172  
A:Molecule type: mRNA  
A:Residues: 1-83, 'A', 85-374, 'RC', 'DIX>  
A:Cross-references: GB:M25631; NID:G538353; PIDN:AAA36741.1; PID:G538354  
R:Dixit, V.M.; Hennessey, S.W.; Grant, G.A.; Rotwein, P.; Frazier, W.A.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5449-5453, 1986

A:Reference number: A05172; MUID:86287276; PMID:3461443  
A:Accession: A05172  
A:Molecule type: mRNA  
A:Residues: 1-83, 'A', 85-374, 'RC', 'DIX>  
A:Cross-references: GB:M14326; NID:G340005; PIDN:AAA61237.1; PID:G553801  
R:Sun, X.; Skorstengaard, K.; Mosher, D.F.  
J. Cell Biol. 118, 693-701, 1992

A:Title: Disulfides modulate RGD-inhibitable cell adhesive activity of thrombospondin.  
A:Reference number: A42927; MUID:92348511; PMID:1379247  
A:Accession: A42927  
A:Molecule type: protein  
A:Residues: 987-1003 <SUN>  
A>Note: Cys-992 is shown to have a free sulfhydryl  
C:Genetics:  
A:Gene: GDB:THB81; TSPI; TSP  
A:Cross-references: GDB:120438; OMIM:188060  
A:Map position: 15q15-15q15  
A:Introns: 23/1  
A>Note: the list of introns may be incomplete  
C:Complex: homotrimer, disulfide linked  
C:Function:  
C:Description: participates in cell migration and adhesion, and in platelet aggregation  
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v  
C:Keywords: beta-hydroxyasparagine; calcium binding; cell adhesion; glycoprotein; trim  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-1170/Product: thrombospondin 1 #status predicted <MAT>  
F:137-375/Domain: von Willebrand factor type C repeat homology <VWC>  
F:1378-429/Domain: thrombospondin type 1 repeat homology <THR1>  
F:434-490/Domain: thrombospondin type 1 repeat homology <THR2>  
F:491-547/Domain: thrombospondin type 1 repeat homology <THR3>  
F:551-586/Domain: EGF homology <EGF1>  
F:650-689/Domain: EGF homology <EGF2>

F:926-928/Region: cell attachment (R-G-D) motif  
F:171-232/Disulfide bonds: #status predicted  
F:248,360,708,1067/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:270,274/Disulfide bonds: interchain #status predicted  
F:610/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
F:1051/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 5.1%; Score 246; DB 1; Length 1170;  
Best Local Similarity 30.3%; Pred. No. 1.3e-09;  
Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVROARLADTANYTCVAKNIVAR-----RRSASAAVIVVVGWSTWTSVCSASC 258  
Db 399 IQGRSCDSLLNRCGSSVQTRTCHIQCDDRKFKQ-----DGGWSHSPWSSCSVTC 451  
QY 259 GRGWQXRSRCTNPAPLNGAFCEGQNVHRTVSSLLVSDGWSWSPWKSACGLDC--- 315  
Db 452 GGVITRILCNPSPPQMGKPCGEARETKACKKDACPFNGWGWSPWDICSVTCGGG 511  
QY 316 THWRSRCDPPAPRNGECQGTDLTRNCTSDLC 350  
Db 512 VOKRSRLCNPTQFGGKDCVGDVTENQCNKQDC 546

RESULT 13  
A39804  
thrombospondin precursor - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A39804  
R:Lawler, J.; Duquette, M.; Ferro, P.  
J. Biol. Chem. 266, 8039-8043, 1991  
A:Title: Cloning and sequencing of chicken thrombospondin.  
A:Reference number: A39804; MUID:91217026; PMID:2022631  
A:Accession: A39804  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1178 <LAW>  
A:Cross-references: GB:M60853; NID:g212763; PID:AA51437.1; PID:g212764  
C:Superfamily: thrombospondin 1; EGF homology; thrombospondin type 1 repeat homology; v  
F:325-383/Domain: von Willebrand factor type C repeat homology <VMC>  
F:386-437/Domain: thrombospondin type 1 repeat homology <THR1>  
F:442-498/Domain: thrombospondin type 1 repeat homology <THR2>  
F:499-555/Domain: thrombospondin type 1 repeat homology <THR3>  
F:658-697/Domain: EGF homology <EGF>

Query Match 5.1%; Score 243.5; DB 1; Length 1178;  
Best Local Similarity 34.4%; Pred. No. 2.1e-09;  
Matches 55; Conservative 15; Mismatches 75; Indels 15; Gaps 4;

QY 210 QARLADTANYTCVAKNIVARRRS-ASAAVIVVNGWSTWTSVCSASCGRGWKRSRS 268  
Db 410 RGRSCDVTRSACTGPHIQTRMCSFKKCDHRIQDGGWSHSPWSSCSVTCGVGNITRIL 469  
QY 269 CTNPAPLNGAFCEGQNVHRTVSSLLVSDGWSWSPWKSACGLDC---THWRSRCD 325  
Db 470 CNGPIPMGGKNCVNGRETEKCEKAPCPNGWGPWSPWSACTVTCGGGIRSRSLCNS 529

QY 326 PAPRNGEECQGTDLDT-----RNCTSDLCVHSAAGP 357  
Db 530 PEQYGGKPCVG---DTKQHDMCNKRCPCPDGCLSNFCFP 566

RESULT 14  
T14160  
transmembrane receptor protein Robo1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T14160  
R:Kidd, T.; Brose, K.; Mitchell, K.J.; Fetter, R.D.; Tessier-Lavigne, M.; Goodman, C.S.;  
Cell 92, 205-215, 1998  
A:Title: Roundabout controls axon crossing of the CNS midline and defines a novel subfam  
A:Reference number: Z17897; MUID:98117249; PMID:9458045

A:Accession: T14160  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-1651 <KID>  
A:Cross-references: EMBL:AF041082; NID:g2811215; PID:g2811216; PID:AA39960.1  
C:Function:  
A:Description: appears to function as the gatekeeper controlling midline crossing  
C:Keywords: transmembrane protein

Query Match 4.6%; Score 220.5; DB 2; Length 1651;  
Best Local Similarity 20.1%; Pred. No. 1.6e-07;  
Matches 205; Conservative 132; Mismatches 377; Indels 305; Gaps 48;

QY 41 DLLPHFLPEDEVIVKKNKPVLLCKAV--PATQIFKCKNGEVRVROVDHVIERS----- 92  
Db 65 DPFPRIHEPSDLIVSKGEPATLNCKAQRPTFTIEWYKGERV-ETDKDDPRSHMLLP 123  
QY 93 -----TDGSSGEPTMEVRINVSROQVEKVFGLLEYWCQCVAMSSGGTTKSQKAYIR 143  
Db 124 SGLSFLRIVHGRKSRPDEGVYI-----CVARNYLGEAVSHNASLE 164  
QY 144 IARLKNFQEPBLAKEVSLQGIIVLPCRPEGIPPAEVEWLNEDLVDSLDPNVYIT-R 202  
Db 165 VAILRDDFRQNPSDVMVAVGEPAVMCEQPGRGHPPTTISWKXKDGSLD---DKDERITIR 221  
QY 203 EHSLVVRQARLADTANYTCVAKNIVARRRSASAAVIVVNGW-----STWTSV-- 253  
Db 222 GGLMITYTRKSDAGKYVCVGNMGERESKVDVTVLERPSPFKRPSNLAVTVDDSAEF 281  
QY 254 -CSASGR-----GWQK-----RSRCTNPAPLNGAFCEGQNVHRT 290  
Db 282 KCEARGDPVPTFGWRKXDCGELPKSRYEIRDHTLTKIKVT--AGDMGSYTCVAENMVGA 339  
QY 291 VSLLSVSDGS-----WSPWKSACGLDCTHWRSRCSBPAP-----RNGEEC----- 335  
Db 340 EASATLTQEPHFHVVKPRDQVVALGRVT--FQCEATGNPQPAIFWRERGSQMLLSYQP 398  
QY 336 -QGTDLDTNCTSDLCVHSAAGPDEVALVGLIAVAVCLVLLVLLVLYVCRKKEGLSD 394  
Db 399 POSRRFSVSTGDLTVTVNQ-RSDVGY-----ICQTL-----N 432  
QY 395 VADSSILTSQFPQVSIKPSKADNPILLTIQPDLSITTTTYQGS-----CPRQDGSPPK- 448  
Db 433 VA-GSIITKAY--LEVTDVIADRPPIVIRGQPNQIVAVDGTTLTSCVATGSPVETILWR 489  
QY 449 -----QLTNCHL-----LSPLG-GRHTLHSSPTSEAEFVSRLSTQNYFRS 490  
Db 490 KQGLVSTQDSRIKQLESGLVQIRYAKLGDGTGRYTCTASTPSGEA----- 534  
QY 491 LPRGTSNMTYGTENFLGRLMIENGTISLIP-----PDAIPRGKIYIYTLHKPEDVRL 546  
Db 535 -----TWSAY-----IEVQFGVPVQPPRPTDNLIPSAP-----SKPEVTDV 572  
QY 547 PLACQTLTASPIVSCGPPGVLLTRPVLAMDHGCEPSPDSWSLRLKKQSCGEGWEQVLH 606  
Db 573 SKNTVTLLWQPNL--SGATPTSIIIEAFSHA---SGSSW-----QTVAENVKTTEFA 620  
QY 607 LGEEAASHLYYQLEASACVYFTE-----QGRFALVG 639  
Db 621 IKGLKFNAYILFLVRAANAYGISDPVKTQDVPPPTQGVHDHKQVORELG----- 674  
QY 640 EALSVAARKLKLFPVACTSLEYNIRVYCLHDTHALKEVVQLEKQLGGOLIOBPRV 699  
Db 675 -----NVVLHNLNPTILSSSSVEV-----HWITVDQSQSIQ-----GYKILYRPS- 714  
QY 700 LHFKDSYHNLRLSIHDVPSSLKWSKLL-----VSYQ--BIPFYHIWNGQRYVLHCTFTL 751  
Db 715 ---GASHGESEWLVEFVTRPTKNSVVIPLDKRGVNYEIKARPFENFQGDSEIKFAKTL 771  
QY 752 ERVSPSTSDLACKLWQVQEGDQGSINFIN-ITKOTRFAEL-----LALSEACVPA 803  
Db 772 EE-RPSAPPRS--VTVSKNDGNGTALIVTWPQPEDTQNGVQOEYKVMCLGNETRYHINK 828

Qy 804 LVGSAFK--IPFLIRQKIISLLPPCRRGA--DWRTLAQKLHLDH----- 846  
Db 829 TVDGSFVVPFLV-PCIRYSVEAASTGAGPGVKSEPOFIQLDHSNPNVSPEDQVSLA 887  
Qy 847 --LSFFASKSPSTA-----MILNLWEARHFPNGNUSQLAAAVAGLGQPDAGLFT 893  
Db 888 QQISDVVKQPAFTAGAACWILLMVFSIWLVRHKKRN--GLSSTVAGIRKVPSTFT 944

RESULT 15  
T30805  
dutt1 protein - mouse  
N;Alternate names: transmembrane receptor protein Robol homolog  
C;Species: Mus musculus (house mouse)  
C;Date: 22-Oct-1999 #sequence\_revision 22-Oct-1999 #text\_change 22-Oct-1999  
C;Accession: T30805  
R;Wu, M.C.; Lowe, N.; Fordham, R.; Rabbitts, P.  
submitted to the EMBL Data Library, July 1998  
A;Description: The mouse homologue of human DUTTL1/H-robol gene: protein sequence and chr  
A;Reference number: Z20879  
A;Accession: T30805  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: mRNA  
A;Residues: 1-1612 <NUM>  
A;Cross-references: EMBL:Y17793; NID:el329712; PID:el329713; PIDN:CAA76850.1  
A;Genetic: experimental source: brain  
A;Gene: dutt1  
A;Map position: 16

Query Match 4.5%; Score 217.5; DB 2; Length 1612;  
Best Local Similarity 19.9%; Pred. No. 2.6e-07;  
Matches 211; Conservative 133; Mismatches 359; Indels 355; Gaps 52;

Qy 41 DLLPHFLVPEPVYVKNKPVLLVCKAV--PATQIFPKCMGEWVRQVDHVIERS----- 92  
Db 26 DFPFRIVEHPSDLIVSGEPATLNCAGRPPTTIIEWYKGERV-ETDKDDPRSHRMLLP 84  
Qy 93 -----TDGSSGEPTMEVRINVRQVKEVFGLEEYWCQCVANSSSGTTKSKAYIR 143  
Db 85 SGLSFLFLRIVHGKSRPDEGVYI-----CVARNYLGEAVSHNASLE 125  
Qy 144 IARLNKFNQEPQELAKEVSLQGVILPCRPPEGIPPAEVEWLRNEDLVPSLDENVYIT-R 202  
Db 126 VAILRDDFRQNSDVNMVAVGEPAVMECQPPRGHPPTISWKDGSPLD--DKDERITIR 182  
Qy 203 EHSLVVRQARLADTANYTCVAKNIVARRBSASAAVIVYNGW-----STWTEWSV-- 253  
Db 183 GGLMLITYTRKSDAGKYVCVGTNMVGERSEVAELTVLERPSEVKRPSNLAVTVDDSAEF 242  
Qy 254 -CSASCGRG-----WQK-----RSRSCTNPAPLNGGAPCEGQNVH 287  
Db 243 KCEA---RGDPVPTVRWKDDGELPKSRYEIRDDHTLKIRKVT--AGDMGSYTCVAENNV 297  
Qy 288 DRTVSSLLSVDSG-----WSPWSKWSAGCLDCTHWSRECSDPAP-----RNGGEFC--- 335  
Db 298 GKAEASATLTVQEPHPHFVVKPRDQVVALGRVT-FCQCATGNPQPAIFWRREGSQNLLFS 356  
Qy 336 ---QOTDLDRNCTSLDLCVHSGASGEDVALYVGLVAVCLVLLVLLVLYVCRKKEGL 391  
Db 357 YQPPQSSSRFSVSQGLDITNVQ-RSDVGYI-----ICOTL----- 392  
Qy 392 DSDVADSSILTSGGQFVSIKPSKADNPHLLTIQPDLSITTTTYQGSIL---CPRQDGSPKPF 448  
Db 393 --NVA-GSIITKAY--LEVTDVADRPPVIRQGPVQVAVDGTLLSLCVATGSPAPTI 447  
Qy 449 -----QLTNGHL---LSPLGG-GRHTLHSSPTSEAEFFVSRSLSTQNY 487  
Db 448 LMRKGVINVSQDSRIKQESVLQIRYAKLGLDTCRYCTASTPSGEA----- 495  
Qy 488 FRSLPRGTSNMTYGTNFNGLGRIMIPNTGISLLIP-----PDALPRGKIYEIYTLHKPE- 542  
Db 496 -----TWSAY-----LEVQETGVQVQPPRPTDNLIPSA-----SKPEV 530

Qy 543 -DV---RLPLAGCOTLLSPIVSCGPPGVLLTRPVLAMDHGCEPSPDSWLSRLKKQCEG 598  
Db 531 TDVSKNTVTILSWQPNLNS-----CATPTSYIIIEAFSHA---SGSSW---QTAAE 573  
Qy 599 SWEQDVLHLGEEAPSHLYYCQLRASACVVFTE-----Q 631  
Db 574 NVKTETFAIKGLKFNALYFLVRAANAYGSDPSQISDPVKTDQVPPTSQGVHKKVQRE 633  
Qy 632 LGRFALVGEALSVAALKRLKLLFAPVACTSLEYNIIRVYCLHDTHTDALKEVVQLEKQLGG 691  
Db 634 LG-----NVVLHLHNPILSSSSEV-----HWTVDQSQYIQ-----GY 668  
Qy 692 QLIQEPRLVLFKDSYHNLRLSIHDVPSLWKSLL-----VSQY--EIPFYHIWNGTQR 743  
Db 669 KILYRPS---GASHGESEWLVEFVRTPTKNSVVIPDLRKGVNVEIKARPPFNEFOGADS 724  
Qy 744 YLHCTFTLERSVSTSLACKLWVQVGEQGSFSINFN-ITKDTREAFEL-----LAL 795  
Db 725 EIKFAKTLEB-AFSAPPRS--VTVSKNDGNGTALLVTWQPPEDTQNGMVQEVKWCIGN 781  
Qy 796 ESEAGVPALVGPSAFKIPFLIRQKIISLDDPPCR-----RGADWRTLAQKLHDS 845  
Db 782 ETKYHINKITVDGSTFSV-----VIPSLVPGIRYSVEVAASTGAGPGVKSEPOFIQLDS 834  
Qy 846 H-----LSPFASKPSPTA-----MILNLWEARHFP--NGNLSQLA 878  
Db 835 HGNVSPEDQVSLAQOISDVVRQPAFIAGIGAACWILLMVFSIWLVRHKKRNGLTSTYA 894  
Qy 879 A-----AVAGLGQPDAGLFTVSE 896  
Db 895 GIRKVPSTFTPTVTYQRGGEAVSSGGRP--GLLINISE 930

Search completed: October 4, 2004, 18:32:42  
Job time : 56 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:11:25 ; Search time 27 Seconds  
(without alignments)

1733.743 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRGLPALLGIVLAWL.....AVAGLQPDAGLFTVSEK 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	280	5.8	1584	1 BAI1 HUMAN	O14514 homo sapien
2	273.5	5.7	1172	1 TSP2 HUMAN	P35442 homo sapien
3	273.5	5.7	1172	1 TSP2 MOUSE	Q03350 mus musculus
4	272.5	5.7	1077	1 SM5A MOUSE	Q62217 mus musculus
5	270.5	5.7	1074	1 SM5A HUMAN	Q13531 homo sapien
6	269	5.6	1170	1 TSP2 BOVIN	Q95116 bos taurus
7	267.5	5.6	1522	1 BAI3 HUMAN	O60242 homo sapien
8	266.5	5.6	1173	1 TSP1 XENLA	P35448 xenopus lae
9	258.5	5.4	1093	1 SM5B MOUSE	O60519 mus musculus
10	258	5.4	1572	1 BAI2 HUMAN	O60241 homo sapien
11	252	5.3	1093	1 SM5B HUMAN	Q9P283 homo sapien
12	248	5.2	1170	1 TSP1 MOUSE	P35441 mus musculus
13	246	5.1	1170	1 TSP1 HUMAN	P07996 homo sapien
14	243.5	5.1	1178	1 TSP2 CHICK	P35440 gallus gall
15	243	5.1	1170	1 TSP1 BOVIN	Q28178 bos taurus
16	208.5	4.4	469	1 PROP_HUMAN	P27918 homo sapien
17	205.5	4.3	470	1 PROP_CAVPO	Q84181 cavia porce
18	202.5	4.2	437	1 PROP_MOUSE	P11680 mus musculus
19	186	3.9	1266	1 NGCA CHICK	Q03696 gallus gall
20	181.5	3.8	867	1 SSPO BOVIN	P98167 bos taurus
21	178.5	3.7	1223	1 AT14 HUMAN	O8WXS8 homo sapien
22	175.5	3.7	1736	1 ZOI1 HUMAN	Q07157 homo sapien
23	172	3.6	630	1 AT5A_RAT	Q9ESP7 rattus norv
24	170.5	3.6	1745	1 ZOI1 MOUSE	P39447 mus musculus
25	169	3.5	934	1 CO6 HUMAN	P13671 homo sapien
26	167.5	3.5	905	1 AT5B MOUSE	P57110 mus musculus
27	166	3.5	837	1 AT5A_HUMAN	O75173 homo sapien
28	162.5	3.4	587	1 CO8B_ONCMY	Q90X85 oncorhynch
29	160.5	3.4	562	1 AT15_MOUSE	P59384 mus musculus
30	160.5	3.4	967	1 AT10 HUMAN	Q9UHI8 homo sapien
31	160.5	3.4	1077	1 AT10 HUMAN	Q9H324 homo sapien
32	159.5	3.3	1224	1 AT16_HUMAN	O8TE57 homo sapien
33	156.5	3.3	860	1 AT56_HUMAN	Q9UKP5 homo sapien

34	156	3.3	950	1 AT15_HUMAN	Q8TE58 homo sapien
35	155	3.2	997	1 AT57_HUMAN	Q9UKP4 homo sapien
36	154.5	3.2	1593	1 AT12_HUMAN	P58397 homo sapien
37	154	3.2	1213	1 AT52_MOUSE	Q8C3W3 m adamcs-2
38	153.5	3.2	1205	1 AT53_HUMAN	O15072 homo sapien
39	152	3.2	967	1 AT51_RAT	Q9WUQ1 rattus norv
40	152	3.2	968	1 AT51_MOUSE	P97857 mus musculus
41	151.5	3.2	890	1 AT58_HUMAN	Q9UP79 homo sapien
42	151.5	3.2	1095	1 AT17_HUMAN	Q8TE56 homo sapien
43	151	3.2	930	1 AT55_HUMAN	Q9UNA0 homo sapien
44	150	3.1	1277	1 CAML_FUGRU	Q98902 fugu rubrip
45	149	3.1	140	1 CO8B_RAT	P55314 rattus norv

## ALIGNMENTS

RESULT 1  
BAIL\_HUMAN  
ID BAIL\_HUMAN STANDARD; PRT; 1584 AA.  
AC O14514;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Brain-specific angiogenesis inhibitor 1 precursor.  
GN BAIL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=98054121; PubMed=9393972;  
RA Nishimori H., Shiratsuchi T., Urano T., Kimura Y., Kiyono K.,  
RA Tatsumi K., Yoshida S., Ono M., Kuwano M., Nakamura Y., Tokino T.;  
RT "A novel brain-specific p53-target gene, BAIL, containing  
RT thrombospondin type 1 repeats inhibits experimental angiogenesis.";  
RL Oncogene 15:2145-2150(1997).  
RN [2]  
RP INTERACTION WITH BAP1.  
RX MEDLINE=98321173; PubMed=9647739;  
RA Shiratsuchi T., Futamura M., Oda K., Nishimori H., Nakamura Y.,  
RA Tokino T.;  
RT "Cloning and characterization of BAI-associated protein 1: a PDZ  
RT domain-containing protein that interacts with BAIL.";  
RL Biochem. Biophys. Res. Commun. 247:597-604(1998).  
CC -!- FUNCTION: LIKELY TO BE A POTENT INHIBITOR OF ANGIOGENESIS IN  
CC BRAIN AND MAY PLAY A SIGNIFICANT ROLE AS A MEDIATOR OF THE P53  
CC SIGNAL IN SUPPRESSION OF GLOBLASTOMA. MAY FUNCTION IN CELL  
CC ADHESION AND SIGNAL TRANSDUCTION IN THE BRAIN.  
CC -!- SUBUNIT: INTERACTS WITH BAP1.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. LIKELY TO BE  
CC CONCENTRATED AT CELL-CELL ADHESION SITES.  
CC -!- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED IN  
CC EXPRESSION IS OBSERVED IN SOME GLIOBLASTOMA CELL LINES AND CANCER  
CC TISSUES.  
CC -!- INDUCTION: By p53.  
CC -!- DOMAIN: THE TSP1 REPEATS INHIBIT IN VIVO ANGIOGENESIS IN RAT  
CC CORNEA INDUCED BY BFGF.  
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.  
CC -!- SIMILARITY: Contains 5 TSP type-1 domains.  
CC -!- SIMILARITY: Contains 1 GPS domain.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AB005297; BAA23647.1; -.

DR PIR; T00026; T00026.  
 DR Genew; HGNC:943; BAIL.  
 DR MTM; 602682; -.  
 DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
 DR GO; GO:0005911; C:intercellular junction; TAS.  
 DR GO; GO:0005911; P:protein binding; TAS.  
 DR GO; GO:0007409; P:axonogenesis; TAS.  
 DR GO; GO:0007422; P:cell adhesion; TAS.  
 DR GO; GO:0007422; P:negative regulation of cell proliferation; TAS.  
 DR GO; GO:0007422; P:peripheral nervous system development; TAS.  
 DR GO; GO:0007422; P:signal transduction; TAS.  
 DR InterPro; IPR000832; GPCR secretin.  
 DR InterPro; IPR001879; hormu receptor.  
 DR InterPro; IPR002023; PKD cys\_rich.  
 DR InterPro; IPR000884; TSP1.  
 DR Pfam; PF00002; 7tm\_2; 1.  
 DR Pfam; PF01825; GPS; 1.  
 DR Pfam; PF02793; HRM; 1.  
 DR Pfam; PF00090; tsp\_1; 5.  
 DR SMART; SM00303; GPS; 1.  
 DR SMART; SM00008; Hormr; 1.  
 DR SMART; SM00209; TSP1; 5.  
 DR PROSITE; PS00221; GPS; 1.  
 DR PROSITE; PS00649; G PROTEIN RECP F2\_1; FALSE NEG.  
 DR PROSITE; PS00650; G PROTEIN RECP F2\_2; FALSE NEG.  
 DR PROSITE; PS00227; G PROTEIN RECP F2\_3; 1.  
 DR PROSITE; PS00261; G PROTEIN RECP F2\_4; 1.  
 DR PROSITE; PS00092; TSP1; 5.  
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;  
 KW Repeat; Cell adhesion.  
 FT SIGNAL 1 30  
 FT CHAIN 31 1584  
 FT DOMAIN 31 948  
 FT TRANSMEM 949 969  
 FT DOMAIN 970 980  
 FT TRANSMEM 981 1001  
 FT DOMAIN 1002 1008  
 FT TRANSMEM 1009 1029  
 FT DOMAIN 1030 1052  
 FT TRANSMEM 1053 1073  
 FT DOMAIN 1074 1093  
 FT TRANSMEM 1094 1114  
 FT DOMAIN 1115 1136  
 FT TRANSMEM 1137 1157  
 FT DOMAIN 1158 1166  
 FT TRANSMEM 1167 1187  
 FT DOMAIN 1188 1584  
 FT TRANSMEM 261 315  
 FT DOMAIN 354 407  
 FT TRANSMEM 409 462  
 FT DOMAIN 467 520  
 FT TRANSMEM 522 575  
 FT DOMAIN 581 938  
 FT TRANSMEM 1411 1422  
 FT DOMAIN 1425 1430  
 FT SITE 231 233  
 FT TRANSMEM 1365 1584  
 FT DOMAIN 1581 1584  
 FT CARBOHYD 64 64  
 FT CARBOHYD 401 401  
 FT CARBOHYD 607 607  
 FT CARBOHYD 692 692  
 FT CARBOHYD 844 844  
 FT CARBOHYD 877 877  
 FT CARBOHYD 881 881  
 FT SEQUENCE 1584 AA; 173531 MW; DEAF28C77874513 CRC64;  
 Query Match 5.8%; Score 280; DB 1; Length 1584;  
 Best Local Similarity 31.8%; Pred. No. 3.8e-13;  
 Matches 74; Conservative 38; Mismatches 93; Indels 28; Gaps 10;  
 QY 124 CQCVAWSSGGTYSQAYTRIALRKNFQEPFLAKEVSLQGVLPFCRPFEGIPPAEVEW 183  
 DR PIR; A47379; TSHUP2.  
 DR EMBL; L12350; AAA03703.1; -.  
 DR EMBL; M81339; -; NOT\_ANNOTATED\_CDS.  
 DR PIR; A47379; TSHUP2.  
 Db 309 CNREACGPAGRTSSRSQSLRSTDARR---REELGDEL---QQGFPPA-PQTGDPAAE-EW 360  
 QY 184 LRNEDLVDPDLDPNVYITRHSLSVVRQARLADTANYTCVAKNIVARRRSASAAVIVYVNG 243  
 Db 361 --SPWSVCSSTCGEGWQTR-----TRFCVSSYSTQCSGLPRLQCNNSAVCPVHG 410  
 QY 244 GWSTWTWVSVCASACGRCGWQKRSCTNPAPLNGAFCGQNVHDRVSSLLV---SVDG 300  
 Db 411 AWDEWSPWSLCSSTCGRGFRDTRTCR--PPQFGNFCGEPKQTKFCNTALCPGRAVDG 468  
 QY 301 SWSFPWSKWSACGLDCT---HWRSGRESDPAPRNGERCQGTDLDRNCTSDLC 350  
 Db 469 NNWEMSSWSACSASCQGRQRTRECNGPS--YGGACQGHWTETDCFLQOC 519  
 RESULT 2  
 TSP2 HUMAN  
 ID TSP2 HUMAN STANDARD; PRT; 1172 AA.  
 AC P35442;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Thrombospondin 2 precursor.  
 GN THBS2 OR TSP2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94010892; PubMed=8406456;  
 RA Labell T.L., Byers P.H.;  
 RT "Sequence and characterization of the complete human thrombospondin 2  
 cDNA: potential regulatory role for the 3' untranslated region.";  
 RL Genomics 17:225-229(1993).  
 RN [2]  
 RP SEQUENCE OF 560-1172 FROM N.A.  
 RX TISSUE=Fibroblast;  
 RA Labell T.L., McGookey Milewicz D.J., Distche C.M., Byers P.H.;  
 RT "Thrombospondin II: partial cDNA sequence, chromosome location, and  
 expression of a second member of the thrombospondin gene family in  
 humans.";  
 RL Genomics 12:421-429(1992).  
 RN [3]  
 RP THROMBOSPONDIN REPEATS DISULFIDE BONDS.  
 RX MEDLINE=21588233; PubMed=11590138;  
 RA Misenheimer T.M., Hahr A.J., Harms A.C., Annis D.S., Mosher D.F.;  
 RT "Disulfide connectivity of recombinant C-terminal region of human  
 thrombospondin 2";  
 RL J. Biol. Chem. 276:45882-45887(2001).  
 CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and  
 cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,  
 laminin and type V collagen.  
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.  
 CC -!- SIMILARITY: Belongs to the thrombospondin family.  
 CC -!- SIMILARITY: Contains 1 WWFC domain.  
 CC -!- SIMILARITY: Contains 3 EGF-like domains.  
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -!- SIMILARITY: Contains 7 TSP type-3 domains.  
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L12350; AAA03703.1; -.  
 CC EMBL; M81339; -; NOT\_ANNOTATED\_CDS.  
 CC PIR; A47379; TSHUP2.

```

DR HSP: P00740; 1EDM.
DR Genew; HGNC:11786; THBS2.
DR GO; GO:0008201; F:heparin binding; TAS.
DR MIM; 188061; -.
DR InterPro; IPR001881; EGF Ca.
DR InterPro; IPR006209; EGF like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR008884; TSP1.
DR InterPro; IPR008885; TSP 1.
DR InterPro; IPR003367; tsp 3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; WFC_C.
DR Pfam; PF00008; EGF; 2.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; vwc; 1.
DR PRINTS; PRO1705; TSP1REPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VWC; 1.
DR PROSITE; PS00022; EGF 1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00026; EGF_3; 2.
DR PROSITE; PS00092; TSP1; 3.
DR PROSITE; PS01208; WFC 1; 1.
DR PROSITE; PS01084; WFC_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1172
FT DOMAIN 19 215
FT DOMAIN 19 232
FT DOMAIN 318 375
FT DOMAIN 381 431
FT DOMAIN 437 492
FT DOMAIN 494 549
FT DOMAIN 549 589
FT DOMAIN 590 647
FT DOMAIN 648 692
FT DOMAIN 725 760
FT DOMAIN 761 783
FT DOMAIN 784 819
FT DOMAIN 820 842
FT DOMAIN 843 880
FT DOMAIN 881 916
FT DOMAIN 917 952
FT DOMAIN 953 1172
FT SITE 928 930
FT DISULFID 266 266
FT DISULFID 270 270
FT DISULFID 393 425
FT DISULFID 397 430
FT DISULFID 408 415
FT DISULFID 449 486
FT DISULFID 453 491
FT DISULFID 464 476
FT DISULFID 506 543
FT DISULFID 510 548
FT DISULFID 521 533
FT DISULFID 553 564
FT DISULFID 558 574
FT DISULFID 577 588
FT DISULFID 594 610
FT DISULFID 601 619
FT DISULFID 622 646
FT DISULFID 652 665
FT DISULFID 659 678
FT DISULFID 680 691
FT DISULFID 707 715
HSP: P00740; 1EDM.
Genew; HGNC:11786; THBS2.
GO; GO:0008201; F:heparin binding; TAS.
MIM; 188061; -.
InterPro; IPR001881; EGF Ca.
InterPro; IPR006209; EGF like.
InterPro; IPR006210; IEGF.
InterPro; IPR008884; TSP1.
InterPro; IPR008885; TSP 1.
InterPro; IPR003367; tsp 3.
InterPro; IPR008859; TSPC.
InterPro; IPR003129; TSPN.
InterPro; IPR001007; WFC_C.
Pfam; PF00008; EGF; 2.
Pfam; PF00090; tsp_1; 3.
Pfam; PF02412; tsp_3; 13.
Pfam; PF05735; TSPC; 1.
Pfam; PF02210; TSPN; 1.
Pfam; PF00093; vwc; 1.
PRINTS; PRO1705; TSP1REPEAT.
SMART; SM00181; EGF; 3.
SMART; SM00209; TSP1; 3.
SMART; SM00210; TSPN; 1.
SMART; SM00214; VWC; 1.
PROSITE; PS00022; EGF 1; FALSE_NEG.
PROSITE; PS01186; EGF_2; 1.
PROSITE; PS00026; EGF_3; 2.
PROSITE; PS00092; TSP1; 3.
PROSITE; PS01208; WFC 1; 1.
PROSITE; PS01084; WFC_2; 1.
Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1172
FT DOMAIN 19 215
FT DOMAIN 19 232
FT DOMAIN 318 375
FT DOMAIN 381 431
FT DOMAIN 437 492
FT DOMAIN 494 549
FT DOMAIN 549 589
FT DOMAIN 590 647
FT DOMAIN 648 692
FT DOMAIN 725 760
FT DOMAIN 761 783
FT DOMAIN 784 819
FT DOMAIN 820 842
FT DOMAIN 843 880
FT DOMAIN 881 916
FT DOMAIN 917 952
FT DOMAIN 953 1172
FT SITE 928 930
FT DISULFID 266 266
FT DISULFID 270 270
FT DISULFID 393 425
FT DISULFID 397 430
FT DISULFID 408 415
FT DISULFID 449 486
FT DISULFID 453 491
FT DISULFID 464 476
FT DISULFID 506 543
FT DISULFID 510 548
FT DISULFID 521 533
FT DISULFID 553 564
FT DISULFID 558 574
FT DISULFID 577 588
FT DISULFID 594 610
FT DISULFID 601 619
FT DISULFID 622 646
FT DISULFID 652 665
FT DISULFID 659 678
FT DISULFID 680 691
FT DISULFID 707 715

```

CC laminin and type V collagen.  
 CC -!- SUBUNIT: Homotrimer; disulfide-linked.  
 CC -!- SIMILARITY: Belongs to the thrombospondin family.  
 CC -!- SIMILARITY: Contains 1 WFCC domain.  
 CC -!- SIMILARITY: Contains 3 EGF-like domains.  
 CC -!- SIMILARITY: Contains 3 TSP type-1 domains.  
 CC -!- SIMILARITY: Contains 7 TSP type-3 domains.  
 CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; L07803; AAA33064.1; -.  
 CC EMBL; M64866; AAA40432.1; -.  
 CC PIR; A42587; A42587.  
 CC HSSP; P00740; 1EDM.  
 CC MGD; MGI:98738; Thbs2.  
 CC InterPro; IPR001881; EGF\_Ca.  
 CC InterPro; IPR006209; EGF like.  
 CC InterPro; IPR006210; IEGF.  
 CC InterPro; IPR000884; TSP1.  
 CC InterPro; IPR008085; TSP 1.  
 CC InterPro; IPR003367; TSP 3.  
 CC InterPro; IPR008859; TSPC.  
 CC InterPro; IPR003129; TSPN.  
 CC InterPro; IPR001007; VWF\_C.  
 CC Pfam; PF00008; EGF; 2.  
 CC Pfam; PF00090; TSP 1; 3.  
 CC Pfam; PF02412; TSP 3; 13.  
 CC Pfam; PF05735; TSPC; 1.  
 CC Pfam; PF02210; TSPN; 1.  
 CC Pfam; PF00093; vwc; 1.  
 CC PRINTS; PR01705; TSP1REPEAT.  
 CC SMART; SM00181; EGF; 3.  
 CC SMART; SM00209; TSP1; 3.  
 CC SMART; SM00210; TSPN; 1.  
 CC SMART; SM00214; VWC; 1.  
 CC PROSITE; PS00022; EGF 1; FALSE\_NEG.  
 CC PROSITE; PS01186; EGF 2; 1.  
 CC PROSITE; PS00026; EGF 3; 2.  
 CC PROSITE; PS00092; TSP1; 3.  
 CC PROSITE; PS01208; VWF\_C; 1.  
 CC PROSITE; PS0184; VWF\_C; 2; 1.  
 CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;  
 CC EGF-like domain; Signal.  
 CC SIGNAL 1 18  
 CC CHAIN 1 1172  
 CC THROMBOSPONDIN 2.  
 CC TSP N-TERMINAL.  
 CC HEPARIN-BINDING (POTENTIAL).  
 CC VWF.  
 CC TSP TYPE-1 1.  
 CC TSP TYPE-1 2.  
 CC TSP TYPE-1 3.  
 CC EGF-LIKE 1.  
 CC EGF-LIKE 2, CALCIUM-BINDING (POTENTIAL).  
 CC EGF-LIKE 3.  
 CC TSP TYPE-3 1.  
 CC TSP TYPE-3 2.  
 CC TSP TYPE-3 3.  
 CC TSP TYPE-3 4.  
 CC TSP TYPE-3 5.  
 CC TSP TYPE-3 6.  
 CC TSP TYPE-3 7.  
 CC C-TERMINAL.  
 CC CELL ATTACHMENT SITE (POTENTIAL).  
 CC SITE 930  
 CC INTERCHAIN (PROBABLE).  
 CC INTERCHAIN (PROBABLE).  
 CC BY SIMILARITY.  
 CC  
 CC DISULFID 397 430 BY SIMILARITY.  
 CC DISULFID 408 415 BY SIMILARITY.  
 CC DISULFID 449 486 BY SIMILARITY.  
 CC DISULFID 453 491 BY SIMILARITY.  
 CC DISULFID 464 476 BY SIMILARITY.  
 CC DISULFID 506 543 BY SIMILARITY.  
 CC DISULFID 510 548 BY SIMILARITY.  
 CC DISULFID 521 533 BY SIMILARITY.  
 CC DISULFID 553 564 BY SIMILARITY.  
 CC DISULFID 558 574 BY SIMILARITY.  
 CC DISULFID 577 588 BY SIMILARITY.  
 CC DISULFID 594 610 BY SIMILARITY.  
 CC DISULFID 601 619 BY SIMILARITY.  
 CC DISULFID 622 646 BY SIMILARITY.  
 CC DISULFID 652 665 BY SIMILARITY.  
 CC DISULFID 659 678 BY SIMILARITY.  
 CC DISULFID 680 691 BY SIMILARITY.  
 CC DISULFID 707 715 BY SIMILARITY.  
 CC DISULFID 720 740 BY SIMILARITY.  
 CC DISULFID 756 776 BY SIMILARITY.  
 CC DISULFID 779 799 BY SIMILARITY.  
 CC DISULFID 815 835 BY SIMILARITY.  
 CC DISULFID 838 858 BY SIMILARITY.  
 CC DISULFID 876 896 BY SIMILARITY.  
 CC DISULFID 912 932 BY SIMILARITY.  
 CC DISULFID 948 1169 BY SIMILARITY.  
 CC CARBOHYD 151 151 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 316 316 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 457 457 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 584 584 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 710 720 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC CARBOHYD 1069 1069 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC SEQUENCE 1172 AA; 129911 MW; 7CE8E4E859822AB CRC64;

Query Match 5.7%; Score 273.5; DB 1; Length 1172;  
 Best Local Similarity 36.1%; Pred. No. 7.9e-13;

Matches 57; Conservative 22; Mismatches 70; Indels 9; Gaps 4;

QY 209 ROARLADTANYTCVAKNIVARRRS-ASAAVIVYVNGSGWSTWTSVCSACGRGWOKRRS 267

Db 403 QRGSCDVTSTNCLGPSIQTRTCSLQKCDTRIRQNGWHSWSSCSVTGCVGNVTRIR 462

QY 268 SCTNPAPLNGGAFCEGQNVHDRTVSLVSDGWSWPSKWSACGLDCT---HWRSRECS 324

Db 463 LCNSPVPQMGKXCKGSGRETKPCORDPCPIDGRNPSWSPNSACTVTCAGGIRERSVCN 522

QY 325 DPAPRNGRECOG--TD---LDTRNCTSLCVHSASGP 357

Db 523 SPEPYGGKDCVGDVTEHQMCKNRCSPIDGCLSNPCFP 560

RESULT 4

SM5A MOUSE

ID SM5A MOUSE STANDARD; PRT; 1077 AA.

AC Q6217;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Semaphorin 5A precursor (Semaphorin F) (Sema F).

GN SEMA5A OR SEMAP OR SEMF.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=NMRI;

RX MEDLINE=96414430; PubMed=8817451;

RA Adams R.H., Betz H., Pueschel A.W.;

RT "A novel class of murine semaphorins with homology to thrombospondin

is differentially expressed during early embryogenesis.";

RL Mech. Dev. 57:33-45(1996).





```

FT DOMAIN 784 839 TSP TYPE-1 5.
FT DOMAIN 841 896 TSP TYPE-1 6.
FT DOMAIN 897 942 TSP TYPE-1 7.
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 277 277 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 367 367 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 437 437 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 536 536 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 717 717 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 933 933 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 56 56 A -> V (IN REF. 2).
FT CONFLICT 149 149 A -> T (IN REF. 2).
FT CONFLICT 382 382 V -> M (IN REF. 2).
FT CONFLICT 494 494 S -> R (IN REF. 2).
SQ SEQUENCE 1074 AA; 120570 MW; EE3DB763CBE29407 CRC64;

Query Match
Best Local Similarity 5.7%; Score 270.5; DB 1; Length 1074;
Matches 50; Conservative 13; Mismatches 52; Indels 3; Gaps 1;

QY 241 VNGGWSWTWESVCSACGRGMRKRSCTNPAPLNGGAFCEGQNVHDTVSSLLVSDG 300
Db 783 VNGANSANTWSQCSRDSCRGIRNRKRVCCNPFKYGGMPCLGSPSLEYQECNTLPCPDG 842
QY 301 SNPSNWKSAAGLDC---THWSRESCDPAPRNGBECQGTDLTRNCTSDLCVHSAS 355
Db 843 VNSCWSPTWKSGATCGGHHYMRTRSCSNPAPAYGGDCLGLHTTEALCNTQPCPSWS 900

RESULT 6
TSP2_BOVIN STANDARD; PRT; 1170 AA.
AC Q95416; Q28180;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 2 precursor (Corticotropin-induced secreted protein) (CISP).
GN THB2 OR TSP2 OR TSP-2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Danik M., Chinn A., Lafeuillade M., Keramidas M., Aquesse-Germon S.,
RA Penhoat A., Chen H., Mosher D., Chambaz E.M., Feige J.J.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-522 FROM N.A.
RX MEDLINE=96331130; PubMed=8698834;
RA Lafeuillade B., Pellerin S., Keramidas M., Danik M., Chambaz E.M.,
RA Feige J.J.;
RP "opposite regulation of thrombospondin-1 and corticotropin-induced
RT secreted protein/thrombospondin-2 expression by adrenocorticotrophic
RT hormone in adrenocortical cells.";
RL J. Cell. Physiol. 167:164-172(1996).
RN [3]
RP SEQUENCE OF 318-831 FROM N.A.
RC TISSUE=Arctic endothelium;
RA Zafar R.S., Moll Y.D., Womack J.F., Walz D.A.;
RT "Cloning and sequencing of bovine thrombospondin stimulatory effect of
RT TGF-beta.";
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin and type V collagen.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.

```

```

CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFCC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X96540; CAA65385.1; -.
CC EMBL; X87620; CAA60952.1; -.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSPC.
CC InterPro; IPR003129; TSPN.
CC InterPro; IPR001007; WVF_C.
CC Pfam; PF00008; EGF; 1.
CC Pfam; PF00090; tsp_1; 3.
CC Pfam; PF02412; tsp_3; 13.
CC Pfam; PF05735; TSPC; 1.
CC Pfam; PF02210; TSPN; 1.
CC Pfam; PF00093; WVC; 1.
CC PRINTS; PRO1705; TSPREPEAT.
CC SMART; SM00181; EGF; 3.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; WVC; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS50026; EGF_3; 2.
CC PROSITE; PS50092; TSP1; 3.
CC PROSITE; PS01208; WVF_1; 1.
CC PROSITE; PS01184; WVF_2; 1.
CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 19 215 THROMBOSPONDIN 2.
FT DOMAIN 19 232 TSP N-TERMINAL.
FT DOMAIN 19 232 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 318 375 WFCC
FT DOMAIN 379 429 TSP TYPE-1 1.
FT DOMAIN 435 490 TSP TYPE-1 2.
FT DOMAIN 492 547 TSP TYPE-1 3.
FT DOMAIN 547 587 EGF-LIKE 1.
FT DOMAIN 588 645 EGF-LIKE 2.
FT DOMAIN 646 690 EGF-LIKE 3.
FT DOMAIN 723 758 TSP TYPE-3 1.
FT DOMAIN 759 781 TSP TYPE-3 2.
FT DOMAIN 782 817 TSP TYPE-3 3.
FT DOMAIN 818 840 TSP TYPE-3 4.
FT DOMAIN 841 878 TSP TYPE-3 5.
FT DOMAIN 879 914 TSP TYPE-3 6.
FT DOMAIN 915 950 TSP TYPE-3 7.
FT DOMAIN 951 1170 C-TERMINAL.
FT SITE 926 928 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 266 266 INTERCHAIN (PROBABLE).
FT DISULFID 270 270 INTERCHAIN (PROBABLE).
FT DISULFID 391 423 BY SIMILARITY.
FT DISULFID 395 428 BY SIMILARITY.
FT DISULFID 406 413 BY SIMILARITY.
FT DISULFID 447 484 BY SIMILARITY.
FT DISULFID 451 489 BY SIMILARITY.

```



```

FT DOMAIN 1045 1098 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1099 1119 6 (POTENTIAL).
FT DOMAIN 1120 1125 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1126 1146 7 (POTENTIAL).
FT DOMAIN 1147 1522 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 159 CUB.
FT DOMAIN 291 343 TSP TYPE-1 1.
FT DOMAIN 345 398 TSP TYPE-1 2.
FT DOMAIN 400 453 TSP TYPE-1 3.
FT DOMAIN 455 508 TSP TYPE-1 4.
FT DOMAIN 816 868 GPS.
FT DOMAIN 942 945 POLY-THR.
FT DOMAIN 1173 1176 POLY-SER.
FT CARBOHYD 51 51 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 82 82 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 105 105 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 241 241 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 337 337 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 625 625 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 779 779 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 812 812 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 828 828 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 937 937 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1522 AA; 171490 MW; D22D0A5D4BB62502 CRC64;

Query Match 5.6%; Score 267.5; DB 1; Length 1522;
Best Local Similarity 37.7%; Pred. No. 3.3e-12;
Matches 55; Conservative 22; Mismatches 54; Indels 15; Gaps 5;

QY 220 TCVA-----KNIVARRRSAAVIVVNGGWSTWTSVCSACGRGWKRSCTNPA 273
Db 317 TCVPSTGTCGSLRSLRVNNTALCPVHGVMWSPWLSCTSCPTCGQRTTRSC--P 374

QY 274 PLMGAGFCQGNVHRTVSLVSDGSWSPWKSACGLDC---THWRSRECSDPAPRN 330
Db 375 PQYGRFCPEGPTHKPCNIALCPVDGQWQWESWQCSVTCSTNGTQORSQCT--AAAH 432

QY 331 GBECCGQGLDTRNCTSDLCVHSASG 356
Db 433 GGECCRGPAWESRECYNPEC--TANG 456

RESULT 8
TSPL_XENLA
ID TSPL_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPL.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RL Submitted (XXX-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIb/beta-3 (by similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWFC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.

```

```

CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L04278; ; NOT_ANNOTATED_CDS.
CC HSSP; P00740; 1EDM.
CC InterPro; IPR001881; EGF_Ca.
CC InterPro; IPR006209; EGF_Like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP_1.
CC InterPro; IPR003367; tsp_3.
CC InterPro; IPR008859; TSP2.
CC InterPro; IPR003129; TSPN.
CC InterPro; IPR001007; VWFC_C.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00090; tsp_1; 3.
CC Pfam; PF02412; tsp_3; 13.
CC Pfam; PF05735; TSPC; 1.
CC Pfam; PF02210; TSPN; 1.
CC Pfam; PF00093; vwc; 1.
CC PRINTS; PR01705; TSPINREPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00209; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS00022; EGF_1; FALSE_NEG.
CC PROSITE; PS01186; EGF_2; 1.
CC PROSITE; PS00026; EGF_3; 2.
CC PROSITE; PS00092; TSP1; 3.
CC PROSITE; PS01208; VWFC_1; 1.
CC PROSITE; PS01184; VWFC_2; 1.
CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
CC EGF-like domain; signal.
CC SIGNAL 1 22
FT CHAIN 23 1173
FT DOMAIN 23 235 THROMBOSPONDIN 1.
FT DOMAIN 23 224 HEPARIN-BINDING (POTENTIAL).
FT DOMAIN 319 376 TSP N-TERMINAL.
FT DOMAIN 382 432 VWFC
FT DOMAIN 438 493 TSP TYPE-1 1.
FT DOMAIN 495 550 TSP TYPE-1 2.
FT DOMAIN 550 590 TSP TYPE-1 3.
FT DOMAIN 591 648 EGF-LIKE 1.
FT DOMAIN 649 693 EGF-LIKE 2.
FT DOMAIN 726 761 EGF-LIKE 3.
FT DOMAIN 762 784 TSP TYPE-3 1.
FT DOMAIN 785 820 TSP TYPE-3 2.
FT DOMAIN 821 843 TSP TYPE-3 3.
FT DOMAIN 844 881 TSP TYPE-3 4.
FT DOMAIN 882 917 TSP TYPE-3 5.
FT DOMAIN 918 953 TSP TYPE-3 6.
FT DOMAIN 954 1173 TSP TYPE-3 7.
FT SITE 929 931 C-TERMINAL.
FT DISULFID 394 426 CELL ATTACHMENT SITE (POTENTIAL).
FT DISULFID 398 431 BY SIMILARITY.
FT DISULFID 409 416 BY SIMILARITY.
FT DISULFID 450 487 BY SIMILARITY.
FT DISULFID 454 492 BY SIMILARITY.
FT DISULFID 465 477 BY SIMILARITY.
FT DISULFID 507 544 BY SIMILARITY.
FT DISULFID 511 549 BY SIMILARITY.
FT DISULFID 522 534 BY SIMILARITY.
FT DISULFID 554 565 BY SIMILARITY.
FT DISULFID 559 575 BY SIMILARITY.
FT DISULFID 578 589 BY SIMILARITY.
FT DISULFID 595 611 BY SIMILARITY.

```

```

FT DISULFID 602 BY SIMILARITY.
FT DISULFID 623 BY SIMILARITY.
FT DISULFID 653 BY SIMILARITY.
FT DISULFID 660 BY SIMILARITY.
FT DISULFID 681 BY SIMILARITY.
FT DISULFID 708 BY SIMILARITY.
FT DISULFID 721 BY SIMILARITY.
FT DISULFID 757 BY SIMILARITY.
FT DISULFID 780 BY SIMILARITY.
FT DISULFID 816 BY SIMILARITY.
FT DISULFID 839 BY SIMILARITY.
FT DISULFID 877 BY SIMILARITY.
FT DISULFID 913 BY SIMILARITY.
FT DISULFID 949 BY SIMILARITY.
FT CARBOHYD 155 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 250 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 363 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 705 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 711 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1173 AA; 130019 MW; A9P036D6516C0F24 CRC64;

Query Match 5.6%; Score 266.5; DB 1; Length 1173;
Best Local Similarity 22.5%; Pred. No. 2.7e-12;
Matches 86; Conservative 54; Mismatches 147; Indels 95; Gaps 14;

QY 11 LIGIVLAAMLRSG-----AQOSTVANVPYGANPDLLPHFLPEPEVDIVVKNKPVLLVC- 65
Db 221 VFGTTLEAILRNKGLSMNSVITLDPNVPNGSPAIRTYIGH-----KTKDLOAVCG 273

QY 66 -KAVPATQIFKNGEVRQVDHVIERTDGSSEPTMEVRINVSQVQEVKVGLEBYWC 124
Db 274 FSCDDLSKLFAMKG-----LRTLVTLKDQVTKETKKNELIA 311

QY 125 QCVAMSSGTTKSKAYIRIARLNKFEQ-----EPLAKEVSLEQGIVLPC 170
Db 312 QIV-----TRTPGVCLHNGVLHKNRDEWTDSCTECTCONSATICRKVSCP---LMFC 361

QY 171 RP-----IPAEVWELRNEDLPBSPDNVYITREHSLVVRQARLADTANY 219
Db 362 TNATIPDGBCCPRCWPSDSADDDSPSWDTPTCS-----VTCGHG-IQQRGSCDSLNN 414

QY 220 TCVAKNIVAR-----RRSAAVIVYVNGWSTWTSVCSACSGRGWQKRSCTN 271
Db 415 PCEGSSVQRTSCIQDCDRFKQ-----DGNWSHSPSSCSVTGSGGQITRIKLS 467

QY 272 PAPLNGGAFCEGQNVHRTVSSLLVSDGSPWSKWSACGLDC---THWRSGRECDPAP 328
Db 468 PVPQLNGKQCEGEGRENKPCQKDPCCPQINGWGFWSLWDTCTVTCGGGMQKRERLCNPXP 527

QY 329 RNGEECCQTDLDTRNCTSDLC 350
Db 528 QYEGKDCIGEPDTSQICNKCQDC 549

RESULT 9
SM5B_MOUSE
ID SM5B_MOUSE STANDARD; PRT; 1093 AA.
AC Q60519;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Semaphorin 5B precursor (Semaphorin G) (Sema G).
GN Sema5B OR SEMA5 OR SEMG.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NMRI;
RX MEDLINE=96414430; PubMed=8817451;

```

```

RA Adams R.H., Betz H., Puschel A.W.;
RT "A novel class of murine semaphorins with homology to thrombospondin
RL is differentially expressed during early embryogenesis.";
CC Mech. Dev. 57:33-45(1996).
CC -!- FUNCTION: May act as positive axonal guidance cues.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: In adult, only detected in brain.
CC -!- DEVELOPMENTAL STAGE: Differentially expressed in embryonic and
CC adult tissues. Its abundance decreases from E10 to birth.
CC -!- SIMILARITY: Belongs to the semaphorin family.
CC -!- SIMILARITY: Contains 1 Sema domain.
CC -!- SIMILARITY: Contains 7 TSP type-1 domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X97818; CAA66398.1; -.
DR MGD; MGI:107555; Sema5b.
DR InterPro; IPR003659; Plexin-like.
DR InterPro; IPR002165; Plexin_repeat.
DR InterPro; IPR001627; Sema.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP_1.
DR Pfam; PF01437; PSI; 1.
DR Pfam; PF01403; Sema; 1.
DR Pfam; PF00090; tsp 1; 5.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00423; PSI; 1.
DR SMART; SM00630; Sema; 1.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS50092; TSP1; 5.
KW Signal, Transmembrane; Repeat; Multigene family; Neurogenesis;
KW Developmental protein; Glycoprotein.
FT SIGNAL 1 19
FT CHAIN 20 1093 SEMAPHORIN 5B.
FT DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 979 999 POTENTIAL.
FT DOMAIN 1000 1093 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 236 518 SEMA.
FT DOMAIN 551 605 TSP TYPE-1 1.
FT DOMAIN 606 662 TSP TYPE-1 2.
FT DOMAIN 664 713 TSP TYPE-1 3.
FT DOMAIN 721 776 TSP TYPE-1 4.
FT DOMAIN 795 850 TSP TYPE-1 5.
FT DOMAIN 852 907 TSP TYPE-1 6.
FT DOMAIN 908 952 TSP TYPE-1 7.
FT CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 944 944 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1093 AA; 120326 MW; 29ESC9B1E8108717 CRC64;

Query Match 5.4%; Score 258.5; DB 1; Length 1093;
Best Local Similarity 22.2%; Pred. No. 1e-11;
Matches 104; Conservative 31; Mismatches 139; Indels 195; Gaps 16;

QY 163 EQGIVLPCRPPEGIPPAEVEWLRNEDLPDNPVYITREHSLVVRQARLADTANYTCV 222
Db 737 EQRPRTCRAP-----LPDP-----HGLQFGKR---TETRTCP 767

```

```
QY 223 AKNIVA-----RRRSAAVIVYVNGGWTWTSVCSACGRGWOKRSCTN 271
Db 768 ADGTGACD TALVEDLLRSGSTSPHTL---NGGWATGWPSSCRDCELGFRVRKTCIN 824
QY 272 PAPLNGGAFCEGQNVHRTVTSLLVSDGWSWPSKWSACGLDC---TWRSSRECSDPAP 328
Db 825 PEPRNGGLPCVGDAAEYQDCNCPACFVRGAWSCWTAWSCQSCGGGHVQRTSCSTSPAP 884
QY 329 RNGGECQGTDLTRNCTSDLCVHSASGPDVAVYGLIAVCLVLLVLLVLYVCRKK 388
Db 885 SPGEDICGLGHTBEALCSTQAC-----PEGWSLW---SEWGVVC-----TE 921
QY 389 EGLDSDVADSSILTSGFPVSIKPKSADNPHLLTIQDLSLTITTYOGSLCPRDGPSPKF 448
Db 922 DGAQSRSCRELLPG-----PGACVGNSSQSRPCPYSEIPV--- 958
QY 449 QLTNGHLLSPGLGGRHTLHSSPTSEAEFVSRSLSTQNYFRSLPRGTSNMITYGTNPLGG 508
Db 959 -----ILPASSVEETTS-----CGGFNL--- 976
QY 509 RLMPNTGISLLIPPAIPRGKIYEIYLTHKPEDVRLPLAGCQTLSPIVSGPPGVLL 568
Db 977 -----IHLIVTG-----VSCFLVSGLL 993
QY 569 TRPVILAMDHCGEPPSDSWSRLKQSCGWSQDVHLHGEAPSHLYY 617
Db 994 TLAVYLSQHCQRQSQES-----TLVH---PATPNHLHY 1024

RESULT 10
BAI2 HUMAN STANDARD; PRT; 1572 AA.
ID BAI2 HUMAN STANDARD; PRT; 1572 AA.
AC G60241;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Brain-specific angiogenesis inhibitor 2 precursor.
GN BAI2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Fetal brain;
RX MEDLINE=98194217; PubMed=9533023;
RA Shiratsuchi T., Nishimori H., Ichise H., Nakamura Y., Tokino T.;
RT "Cloning and characterization of BAI2 and BAI3, novel genes homologous
RT to brain-specific angiogenesis inhibitor 1 (BAI1).";
RL Cytogenet. Cell Genet. 79:103-108(1997).
CC -!- FUNCTION: MIGHT BE INVOLVED IN ANGIOGENESIS INHIBITION.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN BRAIN. ALSO DETECTED IN
CC HEART, THYMUS, SKELETAL MUSCLE, AND DIFFERENT CELL LINES.
CC -!- SIMILARITY: Belongs to family 2 of G-protein coupled receptors.
CC -!- SIMILARITY: Contains 4 TSP type-1 domains.
CC -!- SIMILARITY: Contains 1 GPS domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; AB005298; BAA25362.1; -.
CC PIR; T00027; T00027.
CC Genew; HGNC:944; BAI2.
CC MIM; 602683; -.
DR InterPro; IPR000832; GPCR secretin.
DR InterPro; IPR001879; hormn_receptor.
DR
```

```
DR InterPro; IPR000203; PKD_cys_rich.
DR InterPro; IPR000884; TSP1.
DR Pfam; PF00002; 7tm_2; 1.
DR Pfam; PF01825; GPS; 1.
DR Pfam; PF02793; HRN; 1.
DR Pfam; PF00090; tsp_1; 4.
DR SMART; SM00303; GPS; 1.
DR SMART; SM00008; Hormr; 1.
DR SMART; SM00209; TSP1; 4.
DR PROSITE; PS00221; GPS; 1.
DR PROSITE; PS00649; G_PROTEIN_RECEP_F2_1; FALSE_NEG.
DR PROSITE; PS00650; G_PROTEIN_RECEP_F2_2; FALSE_NEG.
DR PROSITE; PS00227; G_PROTEIN_RECEP_F2_3; 1.
DR PROSITE; PS00261; G_PROTEIN_RECEP_F2_4; 1.
DR PROSITE; PS00092; TSP1; 4.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Repeat.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1572 FT BRAIN-SPECIFIC ANGIOGENESIS INHIBITOR 2.
FT DOMAIN 21 924 FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 925 945 FT 1 (POTENTIAL).
FT DOMAIN 946 953 FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 954 974 FT 2 (POTENTIAL).
FT DOMAIN 975 982 FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 983 1003 FT 3 (POTENTIAL).
FT DOMAIN 1004 1024 FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1025 1045 FT 4 (POTENTIAL).
FT DOMAIN 1046 1066 FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1067 1087 FT 5 (POTENTIAL).
FT DOMAIN 1088 1141 FT CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 1142 1162 FT 6 (POTENTIAL).
FT DOMAIN 1163 1168 FT EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1169 1189 FT 7 (POTENTIAL).
FT DOMAIN 1190 1572 FT CYTOPLASMIC (POTENTIAL).
FT DOMAIN 297 350 FT TSP TYPE-1 1.
FT DOMAIN 352 405 FT TSP TYPE-1 2.
FT DOMAIN 407 460 FT TSP TYPE-1 3.
FT DOMAIN 463 516 FT TSP TYPE-1 4.
FT DOMAIN 859 911 FT GPS.
FT DOMAIN 117 122 FT POLY-GLU.
FT DOMAIN 177 180 FT POLY-ASN.
FT DOMAIN 222 225 FT POLY-THR.
FT DOMAIN 1303 1306 FT POLY-PRO.
FT DOMAIN 1352 1358 FT POLY-GLY.
FT DOMAIN 1413 1418 FT POLY-PRO.
FT CARBOHYD 94 94 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 179 179 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 180 180 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 344 344 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 425 425 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 548 548 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 633 633 FT N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 855 855 FT N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 1572 AA; 171140 MW; A9775645B77BC285 CRC64;

Query Match 5.4%; Score 258; DB 1; Length 1572;
Best Local Similarity 18.8%; Pred. No. 1.9e-11;
Matches 172; Conservative 110; Mismatches 312; Indels 322; Gaps 37;

QY 173 PEGIPPAEVEWLRNEDLVDPISLDPNVY-----ITREHSIVVQARL 213
Db 271 PEEEPKVKQWPRSD-----EPGLYMAQTGPAAEEMSPWSVCSLTCCQGLQVR-TRS 323
QY 214 ATANTYCVAKNIIVARRRSAAVIVYVNGGWTWTSVCSACGRGWOKRSCTNPA 273
Db 324 CVSSPYGLTCSGPLETRERFCNNRSATCPVHGVEEWGWSLCSRCGRGSRNRTCV--P 381
QY 274 PLNGGAFCEGQNVHRTVTSLLVSDGWSWPSKWSACGLDC---TWRSSRECSDPAPR- 329
Db 382 PQHGGKACGPELQTKLCSMAACPVEGQWLENGPWPGPCSTSCANGTQQRSKCSVAGPAW 441
QY 330 -----NGGECQ 336
```

```
Db 442 ATCTGALTDTRECSNLECPATDSKWPANWMSLCSKTCDTGQRRMRMQATGTQGPCE 501
QY 337 GTDLOTRNCTSDLC--VHSASGPEVAL----- 362
Db 502 GTGEEVKPCSEKCPAFHEMCRDEYVYMLTWKKAAGEIYINKCPNASSGSRRLCLSA 561
QY 363 ----VYGLAVAVCL--VLLLVLLVYCRKEGLDSDVADSSILTSGFQVPSIKPSKA 415
Db 562 QGVAVWGLESPFARCSHEVRYLYLSREHLAKQRMAGEMSQVRS-LQELLARITY 620
QY 416 DNPHELLTIQDLSITTTTY-QGSLCRPQDGPSPKFQLT-----NGHLLSPLOG 461
Db 621 SGDLFFSVDLNRVDTFKRATVVSADQVQFFQVSVFMDAENKEKWDDAQVSP--G 678
QY 462 GRHTLHHSPTSEAREFV-----SRLSTQNYFRSLPRG-----TSNMYIGTEN 504
Db 679 SVHELLR-----VVEDFIHLVGDALKAFOSLSLIVTDNLVISIQREBPVSASVSDITFPMRG 732
QY 505 FLG-----GRLMIPNTGISLLIP-----PDAIPRGK----- 530
Db 733 RRGMDWVRHSEDRFLPKEVLSLSGPKPATSGAAGSPGRGPGTVPVPPGCHSHQRL 792
QY 531 -----IYE-IYLTLLHKPEDVRLPLAGCQTLSPISVSCGPPGVLLTRPVIL 574
Db 793 PADPDESSYFVIGAVLYRTLGLILPPP---RPPLAVTSRVMT--VTVRPPTQPPAEPLIT 847
QY 575 A-----MDHCGPSPDSWLRLLKKQCEGSHQDVLHGBEAPSHYYQCLEASACYVT 629
Db 848 VELSVIINGTTDPHCASWDYS-RADASSGSDWTENCQTLTQAAHT-RCQ-----C 896
QY 630 EOLGRFALVGE-----ALSVAAAKRLKLLFAPVACTSLEYNRYVYCLDTHDALKEVVO 684
Db 897 QHLSTFAVLAQPPKDLTLEAGSPVPLVIGCAVSMALLTLAIYA-----AFWRFIK 950
QY 685 LEKQLGGQ-LQBPVRLHFKDSYHNRLSLSDHPSPSLKSLVSYOEIPFYHIWNGTORY 744
Db 951 SERSI-----ILLNFCLSI--LASNI---LILVGSRVLSKGVCTMTAAAF 990
QY 745 LHCTFTLERSVSTSLACKLVW-----LSSFCWVLTBAWQSYLAVIGRMTRLVRKFLCLGMLPALVVA 1039
Db 991 LHFFF-----LSSFCWVLTBAWQSYLAVIGRMTRLVRKFLCLGMLPALVVA 1039
QY 777 FSINENITKOTRFAELLALLESAG-VPALVGPSA-----FKIPPLIROKI-----ISSL 824
Db 1040 VSVGFRTRKGYGTSYVCWLSLEGLLYAFVGAADVIVVNMILGIIVFNKLMARDGISDK 1099
QY 825 DPPCRRGAD---WRTL 837
Db 1100 SKQRAGSERCPWASL 1115

RESULT 11
SM5B HUMAN STANDARD; PRT; 1093 AA.
AC Q9P283;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Semaphorin 5B precursor.
GN SEMA5B OR KIAA1445.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 7:143-150(2000).
```

```
CC -- FUNCTION: May act as positive axonal guidance cues (By
CC similarity).
CC -- SUBCELLULAR LOCATION: Type I membrane protein.
CC -- SIMILARITY: Belongs to the semaphorin family.
CC -- SIMILARITY: Contains 1 Sema domain.
CC -- SIMILARITY: Contains 7 TSP type-1 domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AB040878; BAA95969.1; ALT INIT.
CC Genew; HGNC:10737; SEMA5B.
CC InterPro; IPR003659; Plexin-like.
CC InterPro; IPR002165; Plexin_repeat.
CC InterPro; IPR001627; Sema.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP 1.
CC Pfam; PF01437; PSI; 1.
CC Pfam; PF00090; tsp 1; 5.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00423; PSI; 1.
CC SMART; SM00630; Sema; 1.
CC SMART; SM00209; TSP1; 4.
CC PROSITE; PS50092; TSP1; 5.
CC Signal; Transmembrane; Repeat; Multigene family; Neurogenesis;
CC Developmental protein; Glycoprotein.
CC SIGNAL 1 26 POTENTIAL.
CC CHAIN 27 1093 SEMAPHORIN 5B.
CC DOMAIN 20 978 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 979 999 POTENTIAL.
CC DOMAIN 1000 1093 CYTOPLASMIC (POTENTIAL).
CC DOMAIN 236 518 SEMA.
CC DOMAIN 551 605 TSP TYPE-1 1.
CC DOMAIN 606 662 TSP TYPE-1 2.
CC DOMAIN 664 713 TSP TYPE-1 3.
CC DOMAIN 721 776 TSP TYPE-1 4.
CC DOMAIN 795 850 TSP TYPE-1 5.
CC DOMAIN 852 907 TSP TYPE-1 6.
CC DOMAIN 908 952 TSP TYPE-1 7.
CC CARBOHYD 59 59 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 157 157 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 178 178 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 287 287 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 333 333 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 378 378 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 532 532 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 539 539 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 547 547 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 602 602 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 728 728 N-LINKED (GLCNAC. .) (POTENTIAL).
CC CARBOHYD 944 944 N-LINKED (GLCNAC. .) (POTENTIAL).
CC SEQUENCE 1093 AA; 119866 MW; FIFDFB87CEAF0EF CRC64;
CC
CC Query Match 5.3%; Score 252; DB 1; Length 1093;
CC Best Local Similarity 29.4%; Pred No. 3.3e-11;
CC Matches 68; Conservative 34; Mismatches 85; Indels 44; Gaps 10;
QY 241 VNGGMSTWTSVCSACSGRGWKRSCNTPAPLNGGAFCEQNVHDTVSLVSDG 300
Db 851 VRGASWTSVSPSCASCGGHYQTRTSCTSPSPGEDICLG--LHTEALCATQACE 908
QY 301 SWSPPSKWSACGLDCTHWRSCSDPAPNGGEECGTDLDTNCT-SDL-CVHSASGPE 358
Db 909 GSWPSEWSKCTDDGAQSRSHCELLP--GSSACAGNSSQSRPCFYSIPVILPASSME 966
QY 359 DVALYVG-----LIAVAVCLVL--LLLVLILVYCR--KKEGLDSDVADSSILTSGFQV 408
```



```

Db      967  EATGCAGFNLLHLVATGISCFLSGLLTAVLYLSQCQCQSQSESTL----- 1013
QY      409  SIKPSKADNPHLLTIQPDLSLTTTTCYQSLCPRODGPSP-KFQLTNGHLHSP 458
Db      1014  -VHPATPNHLH-----YKGGGTPKNEKVTPEFKTLNKNLIP 1050

RESULT 12
TSPI_MOUSE
ID TSPI_MOUSE STANDARD; PRT; 1170 AA.
AC P35441;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPI.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92128941; PubMed=1774063;
RA Lawler J., Duquette M., Ferro P., Copeland N.G., Gilbert D.J.,
RA Jenkins N.A.;
RT "Characterization of the murine thrombospondin gene.";
RL Genomics 11:587-600(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92147683; PubMed=1371115;
RA Laherty C.D., O'Rourke K., Wolf F.W., Katz R., Seldin M.F.,
RA Dixit V.M.;
RT "Characterization of mouse thrombospondin 2 sequence and expression
RT during cell growth and development.";
RL J. Biol. Chem. 267:3274-3281(1992).
RN [3]
RP SEQUENCE OF 1-490 FROM N.A.
RX MEDLINE=90375546; PubMed=2398070;
RA Bornstein P., Alfì D., Devarayalu S., Franson P., Li P.;
RT "Characterization of the mouse thrombospondin gene and evaluation of
RT the role of the first intron in human gene expression.";
RL J. Biol. Chem. 265:16691-16698(1990).
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-1ib/beta-3.
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 WFDC domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62450; AAA50611.1; JOINED.
DR EMBL; M62451; AAA50611.1; JOINED.
DR EMBL; M62452; AAA50611.1; JOINED.
DR EMBL; M62453; AAA50611.1; JOINED.
DR EMBL; M62454; AAA50611.1; JOINED.
DR EMBL; M62455; AAA50611.1; JOINED.
DR EMBL; M62456; AAA50611.1; JOINED.
DR EMBL; M62457; AAA50611.1; JOINED.
DR EMBL; M62458; AAA50611.1; JOINED.

```

```

DR EMBL; M62459; AAA50611.1; JOINED.
DR EMBL; M62460; AAA50611.1; JOINED.
DR EMBL; M62461; AAA50611.1; JOINED.
DR EMBL; M62462; AAA50611.1; JOINED.
DR EMBL; M62463; AAA50611.1; JOINED.
DR EMBL; M62464; AAA50611.1; JOINED.
DR EMBL; M62465; AAA50611.1; JOINED.
DR EMBL; M62466; AAA50611.1; JOINED.
DR EMBL; M62467; AAA50611.1; JOINED.
DR EMBL; M62468; AAA50611.1; JOINED.
DR EMBL; M62469; AAA50611.1; JOINED.
DR EMBL; M87276; AAA51063.1;
DR EMBL; J05605; AAA40431.1;
DR EMBL; J05605; AAA40431.1; JOINED.
DR PIR; A40558; AAA40431.1;
DR MGD; MGI:98737; Thbs1.
DR GO; GO:0005615; C:extracellular space; IDA.
DR GO; GO:0016525; P:negative regulation of angiogenesis; IDA.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR006209; EGF_Like.
DR InterPro; IPR006210; IEGF.
DR InterPro; IPR000884; TSPI.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR003367; tsp_3.
DR InterPro; IPR008859; TSPC.
DR InterPro; IPR003129; TSPN.
DR InterPro; IPR001007; WVF_C.
DR Pfam; PF000008; EGF; 2.
DR Pfam; PF00090; tsp_1; 3.
DR Pfam; PF02412; tsp_3; 13.
DR Pfam; PF05735; TSPC; 1.
DR Pfam; PF02210; TSPN; 1.
DR Pfam; PF00093; vwc; 1.
DR PRINTS; PR01705; TSPINREPEAT.
DR SMART; SM00181; EGF; 3.
DR SMART; SM00209; TSP1; 3.
DR SMART; SM00210; TSPN; 1.
DR SMART; SM00214; VNC; 1.
DR PROSITE; PS00022; EGF_1; FALSE_NEG.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS0026; EGF_3; 2.
DR PROSITE; PS0092; TSPI; 3.
DR PROSITE; PS01208; WVF_1; 1.
DR PROSITE; PS0184; WVF_2; 1.
KW Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;
KW EGF-like domain; Signal.
FT SIGNAL 1 18
FT CHAIN 19 1170
FT DOMAIN 19 232
FT DOMAIN 24 221
FT DOMAIN 316 373
FT DOMAIN 379 429
FT DOMAIN 435 490
FT DOMAIN 492 547
FT DOMAIN 549 587
FT DOMAIN 588 645
FT DOMAIN 646 690
FT DOMAIN 723 758
FT DOMAIN 759 781
FT DOMAIN 782 817
FT DOMAIN 818 840
FT DOMAIN 841 878
FT DOMAIN 879 914
FT DOMAIN 915 950
FT DOMAIN 951 1170
FT SITE 926 928
FT DISULFID 270 270
FT DISULFID 274 274
FT DISULFID 391 423
FT DISULFID 395 428
FT DISULFID 406 413
FT DISULFID 447 484
FT DISULFID 451 489

```



DR	PDB; 1LSL; 18-DEC-02.	FT	DISULFID	657	676	BY SIMILARITY.
DR	GlycosuiteDB; P07996; --	FT	DISULFID	678	689	BY SIMILARITY.
DR	Genew; HGNC:11785; THBS1.	FT	DISULFID	705	713	BY SIMILARITY.
DR	MM; 189060; --	FT	DISULFID	718	738	BY SIMILARITY.
DR	GO; GO:0004866; F.endopeptidase inhibitor activity; TAS.	FT	DISULFID	754	774	BY SIMILARITY.
DR	GO; GO:0004871; F.signal transducer activity; TAS.	FT	DISULFID	777	797	BY SIMILARITY.
DR	GO; GO:0007275; P.development; TAS.	FT	DISULFID	813	833	BY SIMILARITY.
DR	InterPro; IPR001881; EGF Ca.	FT	DISULFID	836	856	BY SIMILARITY.
DR	InterPro; IPR006209; EGF-like.	FT	DISULFID	874	894	BY SIMILARITY.
DR	InterPro; IPR006210; EGF.	FT	DISULFID	910	930	BY SIMILARITY.
DR	InterPro; IPR000884; TSP1.	FT	DISULFID	946	1167	BY SIMILARITY.
DR	InterPro; IPR008085; TSP 1.	FT	CARBOHYD	248	248	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	InterPro; IPR003367; tsp 3.	FT	CARBOHYD	360	360	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	InterPro; IPR008859; TSPC.	FT	CARBOHYD	360	360	C-LINKED (MAN. . .)
DR	InterPro; IPR003129; TSPN.	FT	CARBOHYD	385	385	/FTid=CAR_000205.
DR	InterPro; IPR001007; VWF_C.	FT	CARBOHYD	394	394	O-LINKED (FUC. . .).
DR	Pfam; PF00008; EGF; 2.	FT	CARBOHYD	438	438	/FTid=CAR_000206.
DR	Pfam; PF00090; tsp 1; 3.	FT	CARBOHYD	441	441	C-LINKED (MAN. . .)
DR	Pfam; PF02412; tsp 3; 13.	FT	CARBOHYD	441	441	/FTid=CAR_000207.
DR	Pfam; PF05735; TSPC; 1.	FT	CARBOHYD	450	450	C-LINKED (MAN. . .)
DR	Pfam; PF02210; TSPN; 1.	FT	CARBOHYD	498	498	/FTid=CAR_000208.
DR	Pfam; PF00093; vwc; 1.	FT	CARBOHYD	507	507	O-LINKED (FUC. . .).
DR	PRINTS; PR01705; TSP1REPEAT.	FT	CARBOHYD	507	507	/FTid=CAR_000209.
DR	SMART; SM00181; EGF; 3.	FT	CARBOHYD	708	708	C-LINKED (MAN. . .)
DR	SMART; SM00209; TSP1; 3.	FT	CARBOHYD	1067	1067	/FTid=CAR_000210.
DR	SMART; SM00210; TSPN; 1.	FT	CARBOHYD	507	507	O-LINKED (FUC. . .).
DR	SMART; SM00214; VWC; 1.	FT	CARBOHYD	708	708	/FTid=CAR_000211.
DR	PROSITE; PS00022; EGF 1; FALSE_NEG.	FT	CARBOHYD	708	708	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	PROSITE; PS01186; EGF_2; 1.	FT	CARBOHYD	1067	1067	N-LINKED (GLCNAC. . .) (POTENTIAL).
DR	PROSITE; PS01186; EGF_3; 2.	FT	CARBOHYD	507	507	O-LINKED (FUC. . .).
DR	PROSITE; PS00092; TSP1; 3.	FT	CARBOHYD	498	498	/FTid=CAR_000209.
DR	PROSITE; PS01208; VWF 1; 1.	FT	CARBOHYD	708	708	C-LINKED (MAN. . .)
DR	PROSITE; PS01208; VWF 2; 1.	FT	CARBOHYD	1067	1067	/FTid=CAR_000210.
DR	PROSITE; PS01184; VWF 2; 1.	FT	CARBOHYD	507	507	O-LINKED (FUC. . .).
KW	Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;	FT	CARBOHYD	708	708	N-LINKED (GLCNAC. . .) (POTENTIAL).
KW	EGF-like domain; Signal; 3D-structure.	FT	CARBOHYD	1067	1067	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	SIGNAL	1	18			
FT	CHAIN	19	1170			
FT	THROMBOSPONDIN 1.					
FT	HEPARIN-BINDING (POTENTIAL).					
FT	TSP N-TERMINAL.					
FT	VWFC.					
FT	TSP TYPE-1 1.					
FT	TSP TYPE-1 2.					
FT	TSP TYPE-1 3.					
FT	EGF-LIKE 1.					
FT	EGF-LIKE 2.					
FT	EGF-LIKE 3.					
FT	TSP TYPE-3 1.					
FT	TSP TYPE-3 2.					
FT	TSP TYPE-3 3.					
FT	TSP TYPE-3 4.					
FT	TSP TYPE-3 5.					
FT	TSP TYPE-3 6.					
FT	TSP TYPE-3 7.					
FT	C-TERMINAL.					
FT	CELL ATTACHMENT SITE (POTENTIAL).					
FT	INTERCHAIN (PROBABLE).					
FT	INTERCHAIN (PROBABLE).					
FT	DISULFID	270	270			
FT	DISULFID	274	274			
FT	DISULFID	391	423			
FT	DISULFID	395	428			
FT	DISULFID	406	413			
FT	DISULFID	447	484			



CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and cell-to-matrix interactions. Can bind to fibrinogen, fibronectin, laminin, type V collagen and integrins alpha-V/beta-1, alpha-V/beta-3 and alpha-IIb/beta-3. May play a role in dentinogenesis and/or maintenance of dentin and dental pulp.

CC -!- SUBUNIT: Homotrimer; disulfide-linked.

CC -!- TISSUE SPECIFICITY: Odontoblasts.

CC -!- SIMILARITY: Belongs to the thrombospondin family.

CC -!- SIMILARITY: Contains 1 WFCC domain.

CC -!- SIMILARITY: Contains 3 EGF-like domains.

CC -!- SIMILARITY: Contains 3 TSP type-1 domains.

CC -!- SIMILARITY: Contains 7 TSP type-3 domains.

CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; AB005287; BAA21115.1; -.

DR EMBL; X87618; CAA60950.1; -.

DR EMBL; X87619; CAA60951.1; -.

DR PIR; S55501; S55501.

DR GlycoSuiteDB; Q28178; -.

DR InterPro; IPR001881; EGF\_Ca.

DR InterPro; IPR006209; EGF\_like.

DR InterPro; IPR006210; IEGF.

DR InterPro; IPR000884; TSP1.

DR InterPro; IPR008085; TSP\_1.

DR InterPro; IPR003387; TSP\_3.

DR InterPro; IPR008859; TSPCN.

DR InterPro; IPR003129; TSPN.

DR InterPro; IPR001007; VWF\_C.

DR Pfam; PF00008; EGF; 2.

DR Pfam; PF00090; TSP\_1; 3.

DR Pfam; PF02412; TSP\_3; 13.

DR Pfam; PF05735; TSPC; 1.

DR Pfam; PF00210; TSPN; 1.

DR PRINTS; PR01705; TSP1REPEAT.

DR SMART; SM00181; EGF; 3.

DR SMART; SM00209; TSP1; 3.

DR SMART; SM00210; TSPN; 1.

DR SMART; SM00214; VWC; 1.

DR PROSITE; PS00022; EGF\_1; FALSE\_NEG.

DR PROSITE; PS01186; EGF\_2; 1.

DR PROSITE; PS50026; EGF\_3; 2.

DR PROSITE; PS50092; TSP1; 3.

DR PROSITE; PS01208; WFCC\_1; 1.

DR PROSITE; PS50184; WFCC\_2; 1.

KW glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat; EGF-like domain; Signal.

FT SIGNAL 1 18

FT CHAIN 19 1170

FT DOMAIN 19 232

FT DOMAIN 24 221

FT DOMAIN 316 373

FT DOMAIN 379 429

FT DOMAIN 435 490

FT DOMAIN 492 547

FT DOMAIN 549 587

FT DOMAIN 588 645

FT DOMAIN 646 690

FT DOMAIN 723 758

FT DOMAIN 759 781

FT DOMAIN 782 817

FT DOMAIN 818 840

FT DOMAIN 841 878

FT DOMAIN 879 914

FT DOMAIN 915 950

FT THROMBOSPONDIN 1.

FT HEPARIN-BINDING (POTENTIAL).

FT TSP N-TERMINAL.

FT VWF.

FT TSP TYPE-1 1.

FT TSP TYPE-1 2.

FT TSP TYPE-1 3.

FT EGF-LIKE 1.

FT EGF-LIKE 2.

FT EGF-LIKE 3.

FT TSP TYPE-3 1.

FT TSP TYPE-3 2.

FT TSP TYPE-3 3.

FT TSP TYPE-3 4.

FT TSP TYPE-3 5.

FT TSP TYPE-3 6.

FT TSP TYPE-3 7.

FT DOMAIN 951 1170

FT SITE 926 928

FT DISULFID 270 274

FT DISULFID 274 274

FT DISULFID 391 423

FT DISULFID 395 428

FT DISULFID 406 413

FT DISULFID 447 484

FT DISULFID 451 489

FT DISULFID 462 474

FT DISULFID 504 541

FT DISULFID 508 546

FT DISULFID 519 531

FT DISULFID 551 562

FT DISULFID 556 572

FT DISULFID 575 586

FT DISULFID 592 608

FT DISULFID 599 617

FT DISULFID 620 644

FT DISULFID 650 663

FT DISULFID 657 676

FT DISULFID 678 689

FT DISULFID 705 713

FT DISULFID 718 738

FT DISULFID 754 774

FT DISULFID 777 797

FT DISULFID 813 833

FT DISULFID 836 856

FT DISULFID 874 894

FT DISULFID 910 930

FT DISULFID 946 1167

FT CARBOHYD 248 248

FT CARBOHYD 360 360

FT CARBOHYD 708 708

FT CARBOHYD 1067 1067

FT CARBOHYD 1085 1085

FT CONFLICT 805 805

SQ SEQUENCE 1170 AA; 129533 MW; 0DD6ADF3E5FA031A CRC64;

Query Match 5.1%; Score 243; DB 1; Length 1170;

Best Local Similarity 30.3%; Pred. No. 1.8e-10;

Matches 47; Conservative 24; Mismatches 66; Indels 18; Gaps 3;

QY 207 VVQARLADTANYTCVAKNIVAR-----RRSASAAVIVYVNGWMTWESVCSASC 258

Db 399 IQQGRSCDNLNRCGSSVQTRTCHIQCCKRFRKQ-----DGGWHSWSPSSCSVTC 451

QY 259 GRGWQKRSRCTNPAPLNGGAFCEGQNVHRTVSSLLVSDGWSWPSKWSACGLDC--- 315

Db 452 GDGVITRIRLCNSPSPQMNGKPCGKARETKACQKDCSPINGWGPWSPWDICSVTCGGG 511

QY 316 TWRSRSCSDPAPRNGGECQGTDLDRNCTSDLC 350

Db 512 VQKRSRLCNPFPQFGKDCVGDVTENQICNKQDC 546

Search completed: October 4, 2004, 18:29:18

Job time : 31 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 4, 2004, 18:21:41 ; Search time 140 Seconds  
(without alignments)  
2026.078 Million cell updates/sec

Title: US-09-970-944-2

Perfect score: 4787

Sequence: 1 MAVRPLWPLALGIVLAAWL.....AVAGLGQPDAGLFTVSEAEFC 899

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvivirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	4592.5	95.9	898	11 Q8K1S4	Q8K1S4 mus musculus
2	4545.5	95.0	898	11 Q88721	Q88721 rattus norv
3	281.9	58.9	544	4 Q96GP4	Q96GP4 homo sapien
4	2705.5	56.5	931	11 Q08747	Q08747 mus musculus
5	2686	56.1	950	11 Q8CD16	Q8CD16 mus musculus
6	2679.5	56.0	931	13 Q7T2Z5	Q7T2Z5 gallus gall
7	2673.5	55.8	931	4 Q95185	Q95185 homo sapien
8	2570.5	53.7	943	13 Q8UGT4	Q8UGT4 xenopus lae
9	2505.5	52.3	1008	11 Q80Y85	Q80Y85 mus musculus
10	2499	52.2	945	11 Q8K1S3	Q8K1S3 mus musculus
11	2497	52.2	945	11 Q08722	Q08722 rattus norv
12	2493	52.1	945	11 Q9D398	Q9D398 mus musculus
13	2486.5	51.9	934	4 Q8IZJ1	Q8IZJ1 homo sapien
14	2479	51.8	945	4 Q86SN3	Q86SN3 homo sapien
15	2144.5	44.8	956	11 Q8K1S2	Q8K1S2 mus musculus
16	2136	44.6	948	4 Q8WYP7	Q8WYP7 homo sapien

## RESULT 1

ID	Q8K1S4	PRELIMINARY;	PRT;	898 AA.
AC	Q8K1S4;			
DT	01-OCT-2002 (Tremblrel. 22, Created)			
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DE	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Netrin receptor Unc5h1.			
GN	UNC5H1.			
OS	Mus musculus (Mouse)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Engelkamp D.;			
RT	"Cloning of three mouse unc-5 genes and their expression patterns at mid-gestation."			
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AJ487852; CAD32250.1; -.			
DR	MGD; MGI:894682; Unc5h1.			
DR	GO; GO:0004872; F:receptor activity; IEA.			
DR	GO; GO:0007165; P:signal transduction; IEA.			
DR	InterPro; IPR000488; Death.			
DR	InterPro; IPR003599; IG.			
DR	InterPro; IPR007110; IG-like.			
DR	InterPro; IPR000884; TSP1.			
DR	InterPro; IPR008085; TSP_1.			
DR	InterPro; IPR000906; ZUS.			
DR	Pfam; PF00531; Death; 1.			
DR	Pfam; PF00047; ig; 1.			
DR	Pfam; PF00090; tsp; 1; 2.			
DR	Pfam; PF00791; ZUS; 1.			
DR	PRINTS; PR01705; TSP1REPEAT.			
DR	SMART; SM00005; DEATH; 1.			
DR	SMART; SM00409; IG; 1.			
DR	SMART; SM00209; TSP1; 2.			
DR	SMART; SM00218; ZUS; 1.			
DR	PROSITE; PS50835; IG_LIKE; 1.			

Q8iut0 homo sapien  
Q8ot71 mus musculu  
Q8niy2 homo sapien  
Q9h9f3 homo sapien  
Q9nb10 drosophila  
Q9v7b5 drosophila  
Q26262 caenorhabdi  
Q44171 caenorhabdi  
Q9pvd5 petromyzon  
Q8tf26 homo sapien  
Q96sc3 homo sapien  
Q96tw7 homo sapien  
Q8iik1 drosophila  
Q8iv45 homo sapien  
Q8cg21 mus musculu  
Q7tmt3 mus musculu  
Q8cgm0 mus musculu  
Q9uc31 drosophila  
Q9vtt0 drosophila  
Q7yu67 drosophila  
Q8ozf8 mus musculu  
Q8mya8 caenorhabdi  
Q7tt33 mus musculu  
Q8ngw8 homo sapien  
Q8bve5 mus musculu  
Q8cgm1 mus musculu  
Q9tts5 bos taurus  
Q8spm4 bos taurus  
Q8cgb2 mus musculu

## ALIGNMENTS

17	1597.5	33.4	597	4	Q8IUT0
18	1442	30.1	228	11	Q8OT71
19	1223.5	25.6	554	4	Q8NIY2
20	971	20.3	366	4	Q9H9F3
21	961.5	20.1	1072	5	Q9NBL0
22	956.5	20.0	1072	5	Q9V7B5
23	953.5	19.9	947	5	Q26262
24	950.5	19.9	947	5	Q44171
25	694	14.5	199	13	Q9PVD5
26	552.5	11.5	351	4	Q8TF26
27	366	7.6	2673	4	Q96SC3
28	366	7.6	5636	4	Q96RW7
29	311.5	6.5	325	5	Q8I1K1
30	306	6.4	518	4	Q8IV45
31	273.5	5.7	1172	11	Q8CG21
32	273.5	5.7	1172	11	Q7TMT3
33	271	5.7	1582	11	Q8CGM0
34	270.5	5.7	1081	5	Q9UG31
35	269.5	5.6	1083	5	Q9VT70
36	269.5	5.6	1091	5	Q7YU67
37	267	5.6	1522	11	Q8OZP8
38	263.5	5.5	1461	5	Q8MYA8
39	258.5	5.4	1122	11	Q7TT33
40	258	5.4	1573	4	Q8NGW8
41	256	5.3	478	11	Q8BVE5
42	254.5	5.3	1560	11	Q8CGM1
43	249	5.2	685	6	Q9TTS5
44	249	5.2	5146	6	Q8SPM4
45	248	5.2	1171	11	Q8CGB2

DT	01-JUL-1997 (TrEMBLrel. 04, Created)
DT	01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE	Transmembrane receptor UNC5H1.
OS	Rattus norvegicus (Rat).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
NCBI_TaxID=10116;	
[1]	
RN	SEQUENCE FROM N.A.
RP	
RC	TISSUE=Brain, and Ventral spinal cord;
RX	MEDLINE=97271897; PubMed=9126742;
RA	Leonardo E.D., Hincin L., Masu M., Keino-Masu K., Ackerman S.L.,
RA	Tessier-Lavigne M.;
RT	"Vertebrate homologues of C. elegans UNC-5 are candidate netrin
RT	receptors.";
RL	Nature 386:833-838(1997).
DR	EMBL; U87305; AAB57678.1; --
DR	GO; GO:0004872; F:receptor activity; IEA.
DR	GO; GO:0007165; P:signal transduction; IEA.
DR	InterPro; IPR000488; Death.
DR	InterPro; IPR003599; IG.
DR	InterPro; IPR007110; Ig-like.
DR	InterPro; IPR000884; TSP1.
DR	InterPro; IPR008085; TSP1.
DR	InterPro; IPR00906; ZUS.
DR	Pfam; PF00531; death; 1.
DR	Pfam; PF00047; ig; 1.
DR	Pfam; PF00090; tpo.1; 2.
DR	Pfam; PF00791; ZUS; 1.
DR	PRINTS; PR01705; TSP1REPEAT.
DR	SMART; SM00005; DEATH; 1.
DR	SMART; SM00409; IG; 1.
DR	SMART; SM00209; TSP1; 2.
DR	SMART; SM00218; ZUS; 1.
DR	PROSITE; PS50835; IG_LIKE; 1.
DR	PROSITE; PS50092; TSP1; 2.
KW	Receptor.
SQ	SEQUENCE 898 AA; 98840 MW; 7A3CBCB9E7ACA135 CRC64;
Query Match	95.0%; Score 4545.5; DB 11; Length 898;
Best Local Similarity	94.7%; Pred. No. 0;
Matches 852; Conservative 19; Mismatches 26; Indels 3; Gaps 3	
QY	1 MAVRGELMPALLGIVLAALRWGSQAQSATVANPVGANPDLLPHFLVPEDEVYIVKNKP 60
DB	1 MAVRGELMPVLIGVILAAWLKSGSAQSATVANPVGANPDLLPHFLVPEDEVYIVKNKP 60
QY	61 VLLVCXAVPATQIIFKCNGEWVRVDHVIERSTDGSGSEPTMEVINVSQQVEKFVGLE 120
DB	61 VLLVCXAVPATQIIFKCNGEWVRVDHVIERSTDSGLPTMEVINVSQQVEKFVGLE 120
QY	121 EYWCCVAMSSGGTTKSQAYIRIARLKRQFEPLAKEVSLQGIVLCRPPGEGIPPAE 180
DB	121 EYWCCVAMSSGGTTKSQAYIRIARLKRQFEPLAKEVSLQGIVLCRPPGEGIPPAE 180
QY	181 VEWLNRDLVPDLDPNVIITRESLSLVROARLADTANYTCVAKNIVARRRSAAVIVY 240
DB	181 VEWLNRDLVPDLDPNVIITRESLSLVROARLADTANYTCVAKNIVARRRSAAVIVY 240
QY	241 VNGGWSTWTWSVCSASCGRWQKRSCTNPAPLNGGAFCCEGQVHDRTVSSLVSDVG 300
DB	241 VNGGWSTWTWSVCSASCGRWQKRSCTNPAPLNGGAFCCEGQV_ QKTACATLCPVDG 299
QY	301 SHSWPKWSAGCLDCTHWRSHECDPAPRNGBECCGTDLDTNCTSDLCVHSASGPEDV 360
DB	300 SHSWSKWSAGCLDCTHWRSHECDPAPRNGEECRGADLTNCTSDLCVHTASCPEDV 359
QY	361 ALYVGLIAVACLVLILLVLYCRKKEGLDSADVDSILTSGFOPVSIKPSKANPHL 420
DB	360 ALYIGLVAVACLFLILLALGLIYCRKKEGLDSADVDSILTSGFQPVSIKPSKANPHL 419
QY	421 LTIQPDLS -TTTTTYQGSICPRQDGFSPKFQLTNHGLSLPGGRHTLHHSSPTSSEAEFFV 479

PRELIMINARY; PRT; 898 AA.









```
O95185
ID O95185 PRELIMINARY; PRT; 931 AA.
AC O95185;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transmembrane receptor UNC5C.
GN UNC5C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=99000841; PubMed=9782087;
RA Ackerman S.L., Knowles B.B.;
RT "Cloning and mapping of the UNC5C gene to human chromosome 4q21-q23.";
RL Genomics 52:205-208(1998).
DR EMBL; AF055634; AAC67491.1; -.
DR Genew; HGNC:12569; UNC5C.
DR GO; GO:0005042; F-actin receptor activity; TAS.
DR GO; GO:0007411; P-axon guidance; TAS.
DR GO; GO:0007420; P-brain development; TAS.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig-c2.
DR InterPro; IPR000884; TSP1.
DR InterPro; IPR008085; TSP.1.
DR InterPro; IPR000906; ZUS.
DR Pfam; PF00531; death. 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp.1; 2.
DR Pfam; PF00791; ZUS; 1.
DR PRINTS; PR01705; TSP1REPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZUS; 1.
DR PROSITE; PS00835; IG-LIKE; 1.
DR PROSITE; PS0092; TSP1; 2.
KW Immunoglobulin domain; Receptor.
SQ SEQUENCE 931 AA; 103101 MW; EFD71122C98DAB8 CRC64;
```

Query Match 55.8%; Score 2673.5; DB 4; Length 931;  
Best Local Similarity 55.8%; Pred. No. 5.7e-238;  
Matches 507; Conservative 154; Mismatches 221; Indels 31; Gaps 12;

```
QY 9 PALLGIVLAALRGSGAQQS---ATVANPVPGANPDLLPHFLVEPEDVIVTKNKPVLVC 65
Db 26 PAL--ALISASGTGSAQDDDFHELPETFPDPPEPLPHFLIEPEAYIVKKNKPNLYC 83
QY 66 KAVPATQIPKKNGEWVRQVDHVIERTSGSGEPTVEVRINVSROQVEKVFGLBEYWCQ 125
Db 84 KASPATQIPKCNSEWVHQDHDIVDERVDETSGLIVREVSIEISROQVEBELFGEYWCQ 143
QY 126 CVAWSGGTTSKQAYIRIARLNKFEQELAKEVLSLEGIVLPCRPPEGIPPAEVEWLR 195
Db 144 CVAWSAGTTSKRAYIRIARLNKFEQELAKEVLSLEGIVLPCRPPEGIPPAEVEWLR 203
QY 186 NEDLVDPSPDPNVIYTRHSLVVRQARLADTANYTCVAKNIIVARRRSASAAVIVYNGW 245
Db 204 NEDIIDPVEDRNFYITIDHNLIIKQARLSDTANYTCVAKNIIVARRRSASAAVIVYNGW 263
QY 246 STTWESVCSASGRGQKRSRCTNPAPLNGAFCEGQNVHRTVSSLLSVSDGWSWPW 305
Db 264 STTWESVCSNRCRGYQKRTCTNPAPLNGAFCEGQSV-QKIACITTLCPVDGRTWPW 322
QY 306 SKWSACGLDCTHWRRECSPPAPNGGECQGTDLTRNCTSLCVHSASGPDVALYVG 365
Db 323 SKWSTCTECHWRRECSPPAPNGGECQGTDLTRNCTSLCVHSASGPDVALYVG 382
QY 366 L-IAVAVCLVLLLVLLVILVYCRKEGLSDVADSSILTSGFPVSIKPSKADNPHLLTIQ 424
```

```
Db 383 IIVAVIVCLAISVVVALFVYRKNHRDFESDIIDSSALNGGFQPVNKAARQD---LLAVP 439
QY 425 PDL-STTTTYQGSICFRQDGPSPKQLTNGHLLSLPLGGRRHLHES-----SPTSAREEV 479
Db 440 PDLTSAAMYGRPVVALHD-VSDKIPMTNSPILDFPLNLIKIVYNTSGAVSQDDLSERT 498
QY 480 SRLS---TQNYF-----RSLPRGT--SNMTYGTENFLGRGLMIPNTGISLLIPDDA 525
Db 499 SKLSQMTQSLLENFALSLSKNSLARQDPTSCATGFSNLSGLHLLVPNSGVSLIPAGA 558
QY 526 IPRGIYEIYLTLHKPEDVRLPLAGCQTLLSPVSCGPGVLLTRPVILAMDHGCPSPD 585
Db 559 IPQGRVYEMVTVHRKETMRPMDSDQTLLTFVWSCGPGCALLTRPVLLTMHCADPTE 618
QY 586 SWSLRLKQSCGSEWQDVLHGEAPSHLYLCQLEASACYVFTGOLFALVGEALSVA 645
Db 619 DKILLKQAAQOME-DVVVVEENFTTPCYIKLDAEACHILTENLSYALVGHSTTKA 677
QY 646 AAKRLKLLFAPVACTSLEYNIRVYCLDTHDALKEVQLEKQLGGQLIQEPRVLHFKDS 705
Db 678 AAKRLKLAIFGLPCCSSLEYSIRVYCLDQTQALKELHLERQTGGQLLEPKALHFKGS 737
QY 706 YNLRLSIHDVPSSLWKSLLVSYQEIPIYHIIWNGTQRYLHCTFTLERSVSPSTSLACKL 765
Db 738 TNLRLSIHDIAHSLWKSLLAKYQEIPIYHVMSSQRLNCHCTFTLERSVSPSTSLACKL 797
QY 766 WYVWEGDQGSFISNFTKTRFAELLALASEAGVPALVGPSAFKIPFLIRQKLISSLD 825
Db 798 CVRQVEGEGIFQLNCTVSEETGIDPLDLPANTITTVTGPSAFSIPFIRQKLCSSLD 857
QY 826 PCRRGADWRTLAQKLHLDLSHLSFFASKPSPTAMILNLWEARHPFNGLNSQLAAAVAGLG 885
Db 858 AFQTRGHDWRMLAHKLNLDYLNYPATKSSPTGCVILDWEAQNFPGNLSMLAAVLEMG 917
QY 886 QPDAGLFTVSEAE 898
Db 918 RHETVVSAAEGQ 930
```

RESULT 8  
Q8JGT4 PRELIMINARY; PRT; 943 AA.  
AC Q8JGT4;  
DT 01-OCT-2002 (TrEMBLrel. 22, Created)  
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE UNC-5 receptor.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]\_  
RP SEQUENCE FROM N.A.  
RA Anderson R.B., Holt C.E.;  
RT "Expression of UNC-5 in the developing Xenopus visual system.";  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY09459; AAM34486.1; --  
DR GO; GO:0004872; F-actin receptor activity; IEA.  
DR InterPro; IPR000488; Death.  
DR InterPro; IPR003599; Ig.  
DR InterPro; IPR007110; Ig-like.  
DR InterPro; IPR003598; Ig-c2.  
DR InterPro; IPR000884; TSP1.  
DR InterPro; IPR008085; TSP.1.  
DR InterPro; IPR000906; ZUS.  
DR Pfam; PF00531; death. 1.  
DR Pfam; PF00047; ig; 1.  
DR Pfam; PF00090; tsp.1; 2.  
DR Pfam; PF00791; ZUS; 1.  
DR PRINTS; PR01705; TSP1REPEAT.



```
Query Match      52.3%; Score 2505.5; DB 11; Length 1008;
Best Local Similarity 52.7%; Pred. No. 2.3e-222;
Matches 497; Conservative 153; Mismatches 240; Indels 53; Gaps 17;

QY 1 MAVRPGMLVALLGIVLAAM-----LRG--SGAQQSATVANPVPGANPDLLPHFLVPEDEV 53
DB 75 MRARSGVRSALLALLLWDPTPSLAGVDSAGQ---VLPDSYSPAPAEQLPYFLLEPQDA 131

QY 54 YIVKNKPVLLVCKAVPATQIIFKCNQGEWVRQVDHVIERTDSSGSEPTWVRVNSRQOV 113
DB 132 YIVKNKPVLLVCKAVPATQIIFKCNQGEWVRQVDHVIERTDSSGSEPTWVRVNSRQOV 191

QY 114 EKVGLEEYWCQCVANSSGTTKSKAYIRIARLKNFEQPLAKEVSLFQGIIVLPCRP 173
DB 192 EELFGLGDIYWCQCVANSSGTTKSKAYIRIARLKNFEQPLAKEVSLFQGIIVLPCRP 251

QY 174 EGIPPAEVLWLNEDLVDSLDNVTREHSLVVRQARLADTANYTCVAKNIVARRSA 233
DB 252 EGVFAEVLWLNEDLVDSLDNVTREHSLVVRQARLADTANYTCVAKNIVARRSA 311

QY 234 SAAVIVYNGWSTWTEWVSCASCGRWKRSCTNPAPLNGAFCEGQVHRTVSS 293
DB 312 TATVIVYNGWSSWAENSPCNRCGRGWKRTCTNPAPLNGAFCEGQ-AFQKTACT 370

QY 294 LLVSVDSGSPKSKWACGLDCTHWRSCSDPAPNGGEBECQGTDLTNCNCTSDLCV 353
DB 371 TVCPVDGAWTEWKSACSTECARWRSRECMAPPQNGRDCSGTLLDKNCTDGLCVLN 430

QY 354 ASGPEDVALYVGL-TAVAVCLVLLLVILVYCRKKEGLSDVDSS-ILTSQFGPVSTK 411
DB 431 LETSGDVALYAGLVAVFVVAVMVAVGVVYVRNCRDFTDITDSSAALTGFHFVNEK 490

QY 412 PSKADPHLL--TIQDPLSTTT--TYGSLCPRODGPSPKQFQTLNGLHLLPLGGRTHLH 468
DB 491 TARPNNPQLLHSPADPLTASAGIYRGPVYALQDS-ADKIPMTNSPLDPLPSLKIKVTN 549

QY 469 SS-----PTSEAEFVSRSLSTQNYFRS-----LPRGTSNMT 499
DB 550 SSTIGSGGLADGADLLGVLPQTPYGDGP-SRDTHEFLHRSASLGSHLLGLPRDPSSV 608

QY 500 YGTFFNLGGLRLMPTNGISILLIPDAIPRGKIYEIYTLHKPEDVRLPLA-GQOTLLSPI 558
DB 609 SGTFFGLGGLSLPGTGVSLVFNPAIPQKFPYDLVHINKABST-LPLSEGQTVLSPS 667

QY 559 VSCGPGVLLTRVILAMDHCGSPDSLSRLKQSCGSEWQDVHLHGEAPSHLYC 618
DB 668 VTCGPTGLLLCRFVLLTVPHCAEVIAGDMTIFOLKTAHQHWP-EVYTLDETLNTPCYC 726

QY 619 QLEASACYFTEGLGRFALVGEALSVAAAKRLKLLPAPVACTSLEVINRVYCLHDTDA 678
DB 727 QLEAKSCHILLDQLGYVFMGESYSRSVAVRLQALPAPALCTSLYSRLVYCLEDTFVA 786

QY 679 LKEVVOLEKQGLQIQEPRVLHFQDSYHNRLSIHDVPSSLWKSLLYSQYIPIFYHI 738
DB 787 LKEVLELERLTLGGVLYVEEPKPLLFKDSYHNRLSLHDIPHAWRSKLAKYQIPFYHV 846

QY 739 NGTQVILHCTFLTERVSPSTSDACKLWVQVGDGQSFNINITYKDRFAILLALESE 798
DB 847 NGSQRALHCTFLTERHSLASTETFCVCRVQVGEQGIQFQHTTLA-ETPAGSLDLCSEA 905

QY 799 AG--VPALVGPSAFKIPFLRQKIISLDPCCRGADWRTLAQKLHLDLSLSPFASKPSP 856
DB 906 PGNAITQLGPYAFKIPLSLROKICSLDAPNRSRGNDWRLLAQKLSDWRLNYFAIKASP 965

QY 857 TAMILNLWEARHPFNGNLSAAVAGLQDPDAGLFTVSEAC 899
DB 966 TGVILDLWEARQQDDGLNLSALEBWKSEMLVAMATDGC 1008

RESULT 10
Q8KLS3
ID Q8K183
AC Q8K183;
PRELIMINARY; PRT; 945 AA.
```





```

QY 837 LAQKLDHSLSPFRASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQDPAGLFTYSE 896
DB 883 LAQKLSMDRYLNYFATKASPTGVLDDWEARQDDGLNLSLASEENGKSEMLVAMTID 942

QY 897 AEC 899
DB 943 GDC 945

RESULT 12
Q9D398
ID Q9D398 PRELIMINARY; PRT; 945 AA.
AC Q9D398;
DT 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE 6330415E02Rik protein.
GN UNC5H2 OR 6330415E02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gofjebori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL; AK018177; BAB31108.1; -.
DR MGD; MGI:894703; Unc5h2.
DR GO; GO:0007165; P:signal transduction; IEA.
DR InterPro; IPR000488; Death.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR000884; TSPL.
DR InterPro; IPR008085; TSP 1.
DR InterPro; IPR000906; ZU5.
DR Pfam; PF00531; death; 1.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF00090; tsp_1; 2.
DR Pfam; PF00791; ZU5; 1.
DR PRINTS; PR01705; TSPIREPEAT.
DR SMART; SM00005; DEATH; 1.
DR SMART; SM00408; IGC2; 1.
DR SMART; SM00209; TSP1; 2.
DR SMART; SM00218; ZU5; 1.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR PROSITE; PS50092; TSPL; 2.
KW Immunoglobulin domain.
SQ SEQUENCE 945 AA; 103725 MW; 43D33B4524E0CBF2 CRC64;

```

Query Match 52.1%; Score 2493; DB 11; Length 945;  
 Best Local Similarity 52.1%; Pred. No. 3e-221;  
 Matches 497; Conservative 152; Mismatches 241; Indels 64; Gaps 18;

```

QY 1 MAVRPLWPALLGIVLAAW-----LRG--SGAQQSATVANPVPGANPOLLPHFLVEPEDEV 53
DB 1 MRARGVSALLALLCWDPTSPSLAGVDSAGQ---VLPDSVPSAPAPQLPFYELLEPQDA 57

QY 54 YIVKPKPVLVCKAVPATQIFFKCKNGEWRVQDVHVIERTDGSSEGTMEVNRINVSQV 113
DB 58 YIVKPKPVELHCRAPFATQIYFKCKNGEWSQNDHVTQESLDEATGLRVREVGIEVSQV 117

QY 114 EKVFGLIEWCQCVAWSSSGTTKSKAYIRIARLKNFEOBLAKEVSLSEQIVILPCRP 173
DB 118 BELFGLIEDYWCQCVAWSSSGTTKSRAYIRIAYLRKNFDEEFLAKEVLDHEVILLQCRP 177

QY 174 EGIPPAEVEWELRNEDLVDPSPDNVYITREHSLVVRQARLADTANTVCVAKNIARRRRA 233
DB 178 EGVPAEVEWELKNEDVIDPAQDTNFLLIDHNLIIIRQARLSTANTVCVAKNIARRRST 237

QY 234 SAAVIVYNGGMSWTWESVCSACSGRGWKFSRSTNPAPLNGAFCEGQVGHDRTVSS 293
DB 238 AATVIVYNGGWSWAESPCSNRCRGWKRTCTNPAPLNGAFCEGQ-AFOKTACT 296

QY 294 LLVSVGDSWSPKWSACGLDCTHWRSECDPAPRNGEECGQTDLDTNCTSLCV-- 351
DB 297 TVCPVDGAWTEWSKWSACSTECARHWRSECMAPPQNGRGDCSTGLLDSKNCITDGLCVLN 356

QY 352 -----HSASGPEDEVALVYGL-IYAVAVCLVLLLLLVYCRKKEGLDSDVADSS-I 400
DB 357 QRTLNDEPKSHPLETSGDVALYAGLVAVFVVAVLMAEGLVYVRNCRDFDITDSSAA 416

QY 401 LTSGQPVSIKPSKADNPHLL--TIQPDLSITT--TYQSLCPRQDGPSPKFLTNGHLLS 457
DB 417 LGGGHPVNFKTAARPNNPQLLHPSAPPDLTASAGIYRGPVYALQDS-ADKIPMTNSPLLD 475

QY 458 PLGGGRHTLHSS-----PTSEAEFFVRLSTQNYFRS----- 490
DB 476 PLFLSKIKYVNSTTSGSGLADGADLGLVLPFGTYPGDF-SRDTHFLHLSASLSQSHL 534

QY 491 --LPRGTSNMTYGTENFLGGRMLMINTGLISLIPDAIPRGKIYEIYTLTHKPEDVRLPL 548
DB 535 LGLPRDPSSVSGTGLGRLSLPCTGVSLLVPGNAIPQGFYDYLHINKAEAT-LPL 593

QY 549 A-GCQTLLSPVSCGPPGVLLTRPVILAMDHCEPSPDSWSRLKKQSCGSEWQDVHL 607
DB 594 SEGSGTILSPSVTCGPTGLLLCRPVVLTVPHCAEVIAGDWIIFOLKTAHQHWE-EVVTL 652

QY 608 GBEASHLYYCOLEASACVFTFEOIGRPAFVGEALVAAAKRLKILLFAPVACTSEYNI 667
DB 653 DEETLNTPCYCLEAKSCHILLDQLGSYVFMGESYSRAVKRLQLAIFAPALCTSEYSL 712

QY 668 RVYCLDHTDALKEVVQLEKQGLIQBPVILHFKDSYHNLRSLSDHVPSSLWKSLLV 727
DB 713 RYVCLDHTDALKEVVQLEKQGLIQBPVILHFKDSYHNLRSLSDHVPSSLWKSLLV 772

QY 728 SYQEIPFYHIMGTQRYLHCTFTLERSVSPSTBDLACKLWVQVEGQSGFSINFINTKDT 787
DB 773 KYQEIPFYHIMGTQRYLHCTFTLERSVSPSTBDLACKLWVQVEGQSGFSINFINTKDT 831

QY 788 RFAELLALIESEAG--VPALVGPSAFKIPFLIRQKIISLSDPPCCRCGADWETLAKHLDS 845
DB 832 PAGSLDALCSAPGNATITQIGPVAFKIPISIRKICSSLDADPSRGNDRLLAKHLDSMR 891

QY 846 HLSFFASKPSPTAMILNWEARHPNGNLSQAAAAVAGLQDPAGLFTYSEAB 899
DB 892 YLNYFATKASPTGVLDDWEARQDDGLNLSLASEENGKSEMLVAMATDGC 945

```

RESULT 13

Q8IZJ1  
 ID Q8IZJ1 PRELIMINARY; PRT; 934 AA.  
 AC Q8IZJ1;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE	Transmembrane receptor UNC5H2.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=2246081; PubMed=12359238;	
RA	Komatsuzaki K., Dalvin S., Kinane T.B.;	
RT	"Modulation of G(alpha2) signaling by the axonal guidance molecule	
RT	UNC5H2.";	
RL	Biochem. Biophys. Res. Commun. 297:898-905 (2002).	
DR	EMBL; AY126437; AA95701.1; -	
DR	GO:0004872; F:receptor activity; IEA.	
DR	GO:0007165; P:signal transduction; IEA.	
DR	InterPro; IPR000488; Death.	
DR	InterPro; IPR003599; Ig.	
DR	InterPro; IPR007110; Ig-like.	
DR	InterPro; IPR003598; Ig c2.	
DR	InterPro; IPR000884; TSP1.	
DR	InterPro; IPR000805; TSP 1.	
DR	InterPro; IPR000906; ZUS_1.	
DR	Pfam; PF00531; death; 1.	
DR	Pfam; PF00047; Ig; 1.	
DR	Pfam; PF00090; tsp 1; 2.	
DR	Pfam; PF00791; ZUS; 1.	
DR	PRINTS; PR01705; TSP1REPEAT.	
DR	SMART; SM00005; DEATH; 1.	
DR	SMART; SM00409; Ig; 2.	
DR	SMART; SM00408; IGC2; 1.	
DR	SMART; SM00209; TSP1; 2.	
DR	SMART; SM00218; ZUS; 1.	
XW	PROSITE; PS50835; IG LIKE; 1.	
SW	Receptor.	
SW	SEQUENCE 934 AA; 102433 MW; 225B3F506D52B780 CRC64;	
Query Match	51.9%; Score 2486.5; DB 4; Length 934;	
Best Local Similarity	52.2%; Pred. No. 1.2e-220;	
Matches 490;	Conservative 149; Mismatches 255; Indels 45; Gaps 16;	
QY	1 MAVRPGLMAGLIVLAAM--LRGSGAQ--SATVANPVPGANPOLLPHFLVPEPDVIV 56	
DB	1 MGARGGARGALLALLCWDPRLSQAGTDSGEVLPDFFPSAPAEPLPYFLEQPDAYIV 60	
QY	57 KKNPVLLVCKVAPQATIFKCNGEWVRQVDHVIERSDTGSSGEPTVEVINVSRQOQEVK 116	
DB	61 KKNPVLCRAPPATQIYFKCNGEWVSQNDHVTQEGLEATGLRVREVQIEVSRQOQVEEL 120	
QY	117 FGLEEYWCQVAMSSGTTKQKAVIRIARLKNFEQELAKEVLSLEQIVLPCHPPEGI 176	
DB	121 FGLEEYWCQVAMSSAGTTKERRAVRIAYLRKNFDQEPGLKEVPLDHEVLLQCRPPGV 180	
QY	177 PPAEVEWLNEDLVDPSLDPNVYITREHSLVVVRQARLADTANYTCVAKNIVARRRSAAA 236	
DB	181 PVAEWEWLNEDVIDPTQDTNFTLLTIDHLLIQRQLSDTANYTCVAKNIVAKRSTTAT 240	
QY	237 VIVYUNGWSWTEHVSVCASCGRWQKRSCTNPAPINGAFCEGQNVHDTVSSLLV 296	
DB	241 VIVYUNGWSWMAEWSPCNCRGWQKRTCTNPAPINGAFCEGQ-AFQKTACTTIC 299	
QY	297 SVDSGSWSPWSKWSACGLDCTHWSRECDSDPAPNGGECQGTDLDTNCTSDLCVHSAG 356	
DB	300 PVDGAWTEWKSACSTECARHRECMAPPQNGRDCSGTLLDSKNCCTGLCNQMLEA 359	
QY	357 PEDVALYVGL-IAVAVCLVLLVLLVLYCRKKEGLSDVADSS-ILTSGFOPVSIKPSK 414	
DB	360 SGDAALYAGLVVAVFVVAIVAILMAVGVVYVRNCRDFDITDSSAALTGGFHPVNFKTAR 419	
QY	415 ADNEPLL--TTQPDLSITT--TYQSLCPRQDQSPKFOITNGHLLSPILGGGRHTLHHSSP 471	
DB	420 PSNPOLLHSPVPPDLTASAGIRGVPVVALQDS-TDKPMTNSPLDPLPSLKVKVYSST 478	
QY	472 T-----SEAEFVSRSLSTQNY-----FRS-----LPRGTNNMTYGTGTF 503	
DB	479 TSGPGPLADGADLLGLVLPFGTTPSDFARDTHFLHLSASIGSQQLLGLPRDPCGSSVSGTF 538	
QY	504 NFLGGRMLMIPNTGISLLIPDAIPRGKIYEIYLTLHKPEDVRLPLA-GCOTLLSPIVSCG 562	
DB	539 GCLGGRLSIPGTGVSLLVPGALIPQKGFYEMYLINKAEST-LPLSEGTQTVLSPSVTCG 597	
QY	563 PPGVLTFRVILLAMDHCGEFPSPDSRLRLKKSCEGSEWQDVLHLGEEAPSHLYYCOLEA 622	
DB	598 PTGLLLCRPVILTMPHCAEVSARDWIFQLKTAHQHWE-EVVTLDDEILNTFCYQCLEP 656	
QY	623 SACYVTEQIGFALVGEALSVAAAKRLKLLLPAPVACTSLEYNIRVYCLHDTHDALKEV 682	
DB	657 RACHILLDQLTGYVFTGESYSRSYSAVKRLQLAVAFALCTSLSEYSLRVYCLEDTFVALKEV 716	
QY	683 VOLEKQGGQLQIEPRVLHFKDSYHNLRSLHIDVSSLWKSLLVSYOEIPFVHTWNGTQ 742	
DB	717 LELERTLGGYLVVEPKPLMFKDSYHNLRSLHDLPHAWRSKLLAKYQBIPIFYHWSGSO 776	
QY	743 RYLHCTFTLERSVPSSTDLACKLWVQVEGQGSFNFNITKTRFAELLALESEAG-- 800	
DB	777 KALHCTFTLERSLASTELTCKICVQVEGEGQIFQLHTTLA-ETPAGSLDTLCSAPGST 835	
QY	801 VPALVGPSAFKIPFLIRQKIISLDPPCRGRGADWRTLAQKHLDSHLFFAFSKPSPTAMI 860	
DB	836 VTTQLGYPYAFKIPLSIRQKICNSLDAPNSRGNDRMLAQKLSMDRYLYNFATKASPTGVI 895	
QY	861 LNLWEARHPFNGLSOLAAVAGLQGPAGLFTVSEAC 899	
DB	896 LDLWEALQDDGDGDLNSLASALEEMKSEMLVAVATDGDG 934	
RESULT 14		
Q86SN3	PRELIMINARY; PRT; 945 AA.	
ID	Q86SN3	
AC	Q86SN3	
DT	01-JUN-2003 (TREMBLrel. 24, Created)	
DT	01-JUN-2003 (TREMBLrel. 24, Last sequence update)	
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)	
DE	P53-regulated receptor for death and life.	
GN	P53RDL1	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	MEDLINE=22533857; PubMed=12598906;	
RA	Tanikawa C., Matsuda K., Fukuda S., Nakamura Y., Arakawa H.;	
RT	"p53RDL1 regulates of p53-dependent apoptosis.";	
RL	Nat. Cell Biol. 5:216-223 (2003).	
DR	EMBL; AB096256; BAC57998.1; -	
DR	GO:0004872; F:receptor activity; IEA.	
DR	GO:0007145; P:signal transduction; IEA.	
DR	InterPro; IPR000488; Death.	
DR	InterPro; IPR003599; Ig.	
DR	InterPro; IPR007110; Ig-like.	
DR	InterPro; IPR003598; Ig c2.	
DR	InterPro; IPR000884; TSP1.	
DR	InterPro; IPR000805; TSP 1.	
DR	InterPro; IPR000906; ZUS_1.	
DR	Pfam; PF00531; death; 1.	
DR	Pfam; PF00047; Ig; 1.	
DR	Pfam; PF00090; tsp 1; 2.	
DR	Pfam; PF00791; ZUS; 1.	
DR	PRINTS; PR01705; TSP1REPEAT.	
DR	SMART; SM00005; DEATH; 1.	
DR	SMART; SM00409; Ig; 2.	
DR	SMART; SM00408; IGC2; 1.	
DR	SMART; SM00209; TSP1; 2.	
DR	SMART; SM00218; ZUS; 1.	
DR	PROSITE; PS50835; IG LIKE; 1.	



Db 311 VMSWVCSPECBHLRIRECTAPPNGGKFCGLSQESENCTDGLCILDKPLHEIKPQ 370  
QY 352 ----HSASGPEDVALYGLIAVAVCLVLLVLLVLYCRKEGLDSDVADSSILTSGFPQ 407  
Db 371 RNSRRGIEINASDIALYSGL-GAAVAVAVLIVIGVITLYRRSHSDYGVDIVDSSALTGGFQT 429  
QY 408 VSIKPSKADNPHLL--TIQPDLSITTTTQOGLCPRODGPSPKFLQJNGHLLSPLG---- 460  
Db 430 ENFKTVRQNSULLNPNAMPDLTVSTYSGPIC-LQD-PLDKELMTESSEFNPLSDIKVK 487  
QY 461 -----GGRH-----TLHSSPTSEAEFVSRSLSTQNYFR 489  
Db 488 VOSSFMVSLGVSRAEYHGKXHSCTPPHGNRGFTIHPENKT----PYIQNLS----- 537  
QY 490 SLPRGTSNNTYGTNFELGRLMTPNIGISILLIPPDPAIPRGKIYEIYLTLHKPEDVRLPLA 549  
Db 538 SLPTRELTGTGVFGLGRLVMPNGVSVLLIPHGAIPENSWELIYMSINQGEPP-SLOSD 596  
QY 550 GCQTLLSPIVSCGPPGVLLTRPVILAMDHCGEPPSDSWSLRLKQSCGSEWQODVLHGE 609  
Db 597 GSEVLLSPEVTCGPPDMLVTTFALTIPHCADVSSSEHNIHLKKRTQQKWE-EVMSVED 655  
QY 610 EAPSHIYCOLBASACYVTEQIGRALVGEALSVAAAKRLKLLPAPVACTSLEYNIRV 669  
Db 656 ESTS--CYCLDPPFACHVLLDSFGTYALTGEPTDCAVKQKVAVFGCMSCNSLDYLRV 713  
QY 670 YCLHPTHDAKKEVQLEKQGLQIOEPRLVHFKDSYHNRLRSIHDVPSLWKSLLVSY 729  
Db 714 YCVDNTPCAFQEVISDERHQGQJUSEPKLLHFKGNTFSLQVSVLDPPLWRIKPTAC 773  
QY 730 QBIPIFYHNGTQRYLHCTFTLERSVSPSTDLACKLWVQVGDGQSFSGINFINITKDTRF 789  
Db 774 QEVPSRVWSSNRQPLHCAFSLEYRTPTTQJLCKICIRQLKGHEQLQVQTSILESERE 833  
QY 790 AELLALESEAGVPALVGPSAFKIPFLIRQKIISSLDPPCCRRGADWRTLAQKLHLDHLSF 849  
Db 834 TITPFAQEDSTPFAQTGPKAFKIPYSIRQIRICATEDTPNAGKGMQMLAQKNSINENLSY 893  
QY 850 FASKPSPTAMILNWEARHFPNGNLSQLAANAAGLQDPDAGLFTVSEAE 898  
Db 894 FAIQSSPSAVILNWEARHQQDGLDLSLACALEEIRGTHTKLSNITEPQ 942

Search completed: October 4, 2004, 18:31:49  
Job time : 148 secs

**This Page Blank (uspto)**